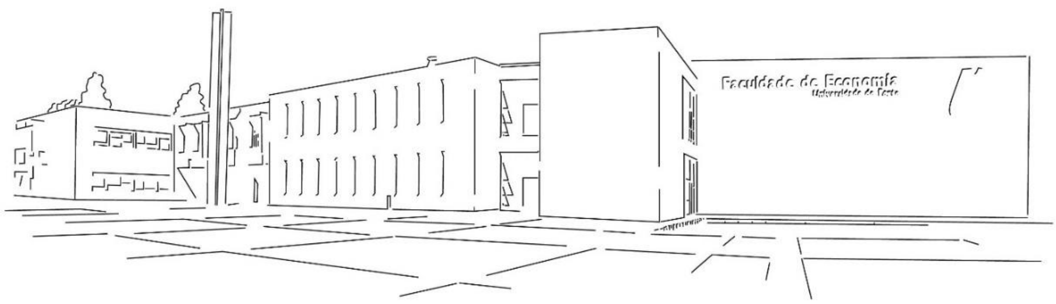
DETERMINANTS OF FIRMS' R&D INVESTMENT DECISIONS: A CROSS COUNTRY STUDY

by

Ana Rita Bento[1]

Master in Finance Dissertation

Advisor: **PhD Júlio Fernando Seara Sequeira da Mota Lobão**



[1] 201208899@fep.up.pt

2016

ABSTRACT

The aim of this study is to investigate the determinants of Research and Development (R&D) investment decisions. Our empirical study uses a logistic model to predict R&D activities, such as the acquisition of existing knowledge, machinery, equipment and other capital goods, training, marketing, design and software development, which can be carried in-house or obtained from third parties.

We focus on internal determinants from financial resources to intangible resources, in order to provide which, determinants affect the overall decision-making process of investing in R&D activities. Our sample is from 2015 listed companies, located in Germany and France, extracted from Thomson Reuters Datastream and R&D intensity, with a total of 556 companies.

Through a logistic regression, the empirical findings suggest that financial autonomy, size and business resources have a high impact on the probability of carrying R&D activity in harmony with the current literature.

The findings of this study can attend as a guide for entrepreneurs and valuable reference in academic research.

Keywords: Innovation, R&D, R&D investment, determinants, France, Germany, Logit

JEL classification codes: G3, O3

BIOGRAPHICAL SKETCH

Ana Rita Amorim Gomes Bento was born August 9, 1987 in Oporto. She received the bachelor's degree in Economics from Lusíada University of Oporto in 2008 and the bachelor's degree in Economic and Business Science from Lusíada University of Vila Nova de Famalicão in 2010.

In 2013, she attended the Executive Course of Taxation and Tax Planning in Católica University of Oporto Business School. In the same year she decided to join the Master in Finance at Faculty of Economics in University of Oporto.

At the end of 2008 she started her professional experience in Santander Totta Bank as Bank Administrative. Between 2010 and 2012 she was a Credit Analyst in BNP Paribas Personal Finance S.A.. During her Master she held the Trainee position in the Financial Department of Nobel Biocare Portugal S.A.. In 2013/2014 and the Financial and Management Controller at Contos da Praça Unip. Lda. At the end of July of 2016 she started a new position at Adidas Financial Shared Service Center.

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor PhD Júlio Fernando Seara Sequeira da Mota Lobão of the Economics Faculty at Oporto University. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it.

I thank my colleagues and friends, but specifically my dearest girlfriends who were always there for me, Rafaela and Cristiana for the stimulating discussions, the advices, the continuous support and for all the experiences we have had for the past three years.

To Ricardo, my boyfriend for his continuous encouragement, support and endless patience through the process of researching and writing this thesis.

To Helga, my cousin that although far gave her precious help in English review. Last but not least, I must express my profound gratitude to my parents and to my brother and cousin for providing me unconditional support and advice throughout my years of study.

This accomplishment would not have been possible without them. Thank you.

Author,

Ana Rita Bento

TABLE OF CONTENTS

FIGURES INDEX

TABLES INDEX

1. Introduction

"*Innovation has long been recognized as an important driver of economic growth*" (Huang, et al. (2015) p. 13). New inventions, such as the electric light bulb, plastic, the telephone, television, the Internet, they all created a big impact in economic growth. In advanced economies, growth is increasingly innovation-driven and R&D is a key factor in developing and applying new technologies that reinforce many of the industries so crucial to economies' competitiveness. The technological factor became a critical source of competitive advantage (Urbancová (2013)).

The fast-paced high-tech boom is what finally gave broad currency to Schumpeter's idea of *Creative Destruction (1934)*. Like many powerful ideas, it is simple: a market economy will incessantly revitalize itself from within by scrapping old and failing businesses and then reallocating resources to newer, more productive ones. It became a well-known expression, especially in Silicon Valley, where companies are constantly remaking themselves and new businesses were flaring up and flaming out.

R&D has a number of characteristics that make it different from ordinary investment. Wages and salaries of highly educated scientists and engineers represent most of R&D spending. They create knowledge for firms, intangible assets, from which profits in future years will be generated. When such workers leave or are fired, part of the resource base of the firm disappears.

As an investment R&D has also an important characteristic, the degree of uncertainty. Uncertainty tends to be greater at the beginning of a research program or project, when associated with R&D output, which implies that an optimal R&D strategy has an options-like character and should not really be analysed in a static framework. Between conception and commercialization; R&D projects requires a long times and firms tend to smooth their spending in R&D, essentially because of the degree of uncertainty (Hall (2010)).

The external technology may discourage own in-house R&D investment and there are also arguments indicating that external R&D may stimulate rather than substitute own R&D activities (Arora and Gambardella (1990), Braga and Willmore (1991), Freeman

(1991), Siddharthan (1992)). Therefore, it was imperative for us to study the determinants of R&D investment decisions.

In order to contribute to the literature, we decided to use firm internal determinants, using data not only from manufacture companies but also from companies in sectors with R&D intensity. The most relevant result that we obtained was that financial autonomy, size and business resources are all statistically significant with a positive relation to R&D investment decision, and a predictable power over 50%.

Our study suggests financial autonomy, size and business resources are determinants of R&D investment not only for manufacture companies but also for other sectors such as software and computer services.

The paper is organised as follows. Section 2 contains the literature review of relevant materials related to topic and other literature related to problem as well as the main studies. Section 3 includes the research hypothesis; financial resources, physical resources and intangible resources and in the last point the model. In section 4, we will address data and methodology, description of approach to collecting data, description of theoretical framework for analysis, descriptive statistics and the model formulated. Estimation results are given in Section 5. Some concluding remarks are in section 6.

2. Literature review

To provide the literature review on the field of R&D investment decisions, firstly we will present the main concepts in section 2.1, then the founded theories in section 2.2 and finally the main studies which are the most important section of this chapter for our study.

The literature on research and development (R&D) investment has generally been based on the innovation theory of Schumpeter (1942), which argues that innovation is the foundation of economic long-term growth and success.

2.1. Concepts

Already in 1776, Adam Smith identified the effect that innovation has in economic development. However, it was J. Schumpeter who had a pioneer role in introducing innovation into economic studies. Schumpeter (1934) considered that innovation is the strategic stimulus to economic development, defined as the commercial or industrial application of something new - a new product, process or method of production, a new market or source of supply, a new form of commercial, business or financial organization.

The Oslo Manual Guidelines for collecting and interpreting innovation data go beyond the definition of Schumpeter and consider, "*an innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organization method in business practices, workplace organization or external relations*" (OECD (2005) p. 46). "*Innovation activities are all scientific, technological, organizational, financial and commercial steps which actually, or are intended to, lead to the implementation of innovations*." (OECD (2005) p. 47).

R&D plays a critical role in the innovation process and is a key factor in developing new competitive advantages. Hunt and Morgan (1996) alleged that a new product or process can be a source of market advantage for the innovator, i.e. innovation plays a key role in resource-advantage theory. Some companies rely on R&D for growth through new products, others use R&D to stimulate incremental improvements; some companies conduct no R&D and view it as an unnecessary expense.

2.2. Founded theories

Schumpeter (1939) wrote that technological innovations are the driving force of economic growth. He argued that quasi-monopolistic profits gained by the initiators led to a swarming effect where imitators jumped on the entrepreneurial movement which generated a wave of new investments. He considered the notion of radical technological innovations as a major factor in recurrent structural crises, comparing technological innovations with a series of explosions rather than with a gentle, though incessant, transformation. He found that exogenous factors such as wars were very important.

Schumpeter (1939) based his conclusion on three grounds; First, he argued that innovations are not distributed over the whole economic system at random, but tend to concentrate in certain sectors and their surroundings. Secondly, he argued that the diffusion process is also inherently an uneven one because innovations do not remain isolated events, and are not evenly distributed in time, on the contrary. Finally, the characteristics of points one and two imply that the disturbances the diffusion process engenders are enough to disrupt the existing system and enforce a distinct process of adaptation.

What stimulates the capitalist engine comes from new consumers' goods, new methods of production or transportation, new markets, new forms of industrial organization that capitalist enterprise creates. The opening up of new markets, foreign or domestic, and the organizational development illustrates the same process of industrial mutation that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one. This process was called the *Creative Destruction* and consider this the essential fact about capitalism by Schumpeter (1942).

Schumpeter (1934) established two different patterns; in the first, he believed that innovation was envisaged as a more routinized process with large firms, and in the second Schumpeter (1942) became aware of the rise of in-house corporate R&D in large firms, thinking towards the role of large oligopolistic firms as the key agent for innovation.

However, it was Solow (1956) who attributed the economic growth to technological innovation. He claims that because capital is produced based on known technology, and because technology is improving, new capital is more valuable than old capital. He characterized technological improvements or developments as the improvements in diverse business processes or products and maintained that these improvements turn out to be the innovations which facilitate economic growth. Solow also hypothesized that new technologies drive economic growth to “technical change”.

Nevertheless, it was Romer (1994) who considered that endogenous growth is created by accumulating technology. Romer also determined the direct link between human capital and technological growth. Romer believed that innovation has no external cause. When a company increases its investment in capital, it is not only increasing its own production levels, but also those adjacent companies. He considered knowledge essential to productivity.

2.3. Main studies

There are several empirical literatures about R&D, although most of them study the determinants of R&D expenditures or R&D intensity. R&D expenses are a line item from many companies’ income statements that reflects the amount of money already spent to develop new products and services each year. There is a big difference between R&D expenditures and R&D investment studies. The R&D investment literature evaluate which firms do or do not invest in R&D.

When companies have low R&D expenditures, public support (subsidies) arises finding a way to allocate more resources to R&D. In companies that have already R&D expenditures, the other way to incentivize their investment is to influence the determinants of R&D expenditures of the firms. Encouraging the firms on their first R&D investment provides more R&D activities, thus, increasing probability for innovation. Therefore, identifying the determinants of investment probability in R&D is very important. At this point, we will only mention the studies relevant for this paper.

Cohen, *et al.* (1987) was probably the first attempt to investigate the determinants of the probability of R&D investment decision. Their sample included R&D expenditures from 2.494 business units in 244 manufacturing lines of business operated by 345 firms in the

USA. Their Tobit regression yielded results that lead to the conclusion that the size of the business unit rather than the size of the firm as a whole affects the likelihood of performing R&D. Cohen and Klepper (1996) found in two simple regressions of business unit R&D against, respectively, business unit and overall firm size for each of 75 industries, that business unit size alone explained an average of 65% of the variance in business unit R&D, and the relationship was typically proportionate. In contrast, overall firm size explained an average of only 15%. Moreover, the coefficient of business unit size was positive and significant for almost 90% of the 75 industries, whereas the coefficient of overall firm size was rarely significant, and, indeed, was actually negative for 26 of the 75 industries. Both Cohen, *et al.* (1987) and Cohen and Klepper (1996) found that controlling for business unit size and overall firm size exercised no independent influence on business unit R&D. These results suggest that it is the size of the business unit (or its correlates) rather than the size of the firm as a whole (or its correlates) that accounts for the close relationship between firm size and R&D.

Braga and Willmore (1991) decided to measure the effect of selected variables on the decision of the firm to engage in five technological activities, including imports of technology, new product development, research and development, and quality control. Their data comes from 4.342 firms in Brazil for 1981 and they used discrete dependent variables on a logistic regression model. They used a binary dependent variable which indicates that firms invest or not in R&D. They concluded that firm size (mean value added), exports, foreign technology and diversification of output have a positive and highly significant effect on the probability of R&D investment. The firm size coefficient, though significant in a statistical sense, is very small. They used other independent variables such as foreign control, foreign share of industry output, state control, operating profit of the firm and effective protection; they are all insignificant.

Galende Del Canto and Suárez González (1999) investigated the determinants of the firm's decision to carry out R&D activities. They used a sample of 100 firms under the Spanish law, to disclose their annual accounts for 1992 in the commercial registry of Castilla y León. Empirical findings suggest that intangible factors are the main determinants of the probability of a firm carrying out internal R&D. This is a significant finding because the intangible factors are often thought of as the most important from a strategic point of view, and our study confirms their importance in the strategic decision of investing in R&D. Finally, financial resources are shown to be hardly relevant in the model, and contrary to what was expected, a greater volume of firm equity resources is negatively related to the probability of carrying out R&D activities.

Griffith, *et al.* (2006) estimated the determinants of whether a firm engaged in R&D continuously, as dummy dependent variable over the period 1998–2000 at the firm level for the four major European countries—France, Germany, Spain, and the UK—using data from the third wave of the internationally harmonized community innovation surveys, providing information for the period of 1998-2000. They used a discrete Probit model and independent variables such as international competition, funding and firm size (set according to the firm's number of employees in 1998). All variables have positive and significant effect except for strategic measures in Spain, national funding in France and process innovation also in France which have negative impact. These variables, respectively, have negative and positive but insignificant effect.

Costa-Campi, *et al.* (2014) used a panel data for the period of 2004-2010 from Spanish firms to estimate R&D intensity and R&D decision. They used Tobit regression and Probit regression to estimate for the decision to spend on R&D or not. The results showed that size (number of employees) has opposite effects. It is more likely that bigger energy firms engage in R&D activities but once they carry out R&D activities, smaller companies devote more resources to R&D (in relative terms). And their results led them to reject the Schumpeterian hypothesis and to favour evidence on the advantage of small firms in R&D intensity in the energy industry. They also found that younger firms engage more than older firms in the decision to perform R&D and they are also more likely to devote more resources to R&D activities.

Lai, *et al.* (2015) investigated the decision factors regarding R&D investment activities in Taiwan, Japan and Korea. They used several hypotheses considering internal resources, from financial resources, to physical resources and intangible resources. The model used was logistic regression. They conclude that among the seven major factors, an enterprise's internal and intangible resources are the most important factors affecting the probability of implementing internal investment activities in R&D. This study also confirms the importance of the strategic decisions behind an enterprise's investment activities in R&D.

They suggest that lower financial autonomy can lead to more active investments in R&D and promote the overall growth of a Japanese company. As for intangible resources, all countries indicate a positive significant impact, except for Korea. Such arguments are similar to other studies. By understanding the major factors, they realise that an enterprise can be less susceptible to the influence of uncertain factors when making decisions relating to investing capital in R&D.

Limanlı (2015) investigated the determinants of R&D investment using firm's micro-data level in Turkey for 2008-2013. They used R&D as binary dependent variable in a generalized linear mixed model for complex survey design. Their estimation results verify Schumpeterian hypothesis that bigger size (sales) increase the probability of investment in R&D, but after some point, if firm size continues to increase, the probability of R&D investment begins to decrease. They found also that sales, government subsidies, share of foreign ownership, competition incentive, scale of enterprise, domestic and foreign trade shares are significantly important factors for investment in R&D.

Most of these studies use exclusively manufacturing or any type of grouping companies on their data which could lead to biased results. Nevertheless, the majority has in common that size or its correlates do matter for R&D investment decisions. These studies reveal that internal factors, such as intangible factors, exports, international competition, foreign technology, intangible resources, sales, subsidies, funding and diversification have statistical significance to predict R&D activities. However, several of these factors have always contradictory results, presenting different signs regarding R&D activities, probably having as a factor the way the variables are calculated.

There is a gap in this matter, as far as we know; recently there has been no prediction to R&D investment using European countries as a whole. We will consider firm internal determinants, and how these determinants could influence R&D activities. Appendix 1 summarizes the more relevant and most recent studies.

After having discussed the theoretical background regarding the determinants of R&D investment decisions, in section 3, we will present the hypotheses development and research design.

3. Hypotheses development and research design

Below we will focus in formulating the research hypotheses taking into consideration the literature review. In order to answer the research question, we will follow the previous studies focusing and distinguishing between tangible and intangible resources. This section is divided into research hypotheses, section 3.1 and empirical model, section 3.2.

3.1. Research hypotheses

Schumpeter (1934, 1942) proposes two factors is his studies, the size of the company and the market concentration, as direct determinants of innovation, although he did not consider other internal factors that also determine innovative activity or how the innovative process occurs inside the firm.

Galende Del Canto and Suárez González (1999) clearly confirm the relevant role of resources and capabilities in R&D. They stated that each firm, as a function of its history and past success, has different resources available and this characteristic structure of resources is shown to be a relevant factor when explaining the differences in the investment decisions of a firm. They found that the most relevant group of resources is the intangible (Human and Commercial) ones supported by empirical findings that also suggest that intangible factors are the main determinants of the probability that a firm carries out internal R&D.

The following sections describe the research hypothesis divided into the main topics, financial resources, physical resources and intangible resources.

3.1.1. Financial resources

The available funds can affect the carrying out of R&D activities. Financial resources of a company represent the monetary means held by the same organization or the ability to obtain them in the financial markets or from donors, and can be used to finance the current activity or finance new investments.

3.1.1.1. Financial autonomy

Lai, *et al.* (2015) stated that an enterprise's financial situation is vital to the short and long-term development of a company. Kim and Park (2012) argued that more financial slack would lead to more R&D activities.

Galende and Fuente (2003) said that the accessibility of internal resources could allow greater risk of R&D activities. On the contrary, debt financing is weighted in consideration to opportunistic R&D investment decisions, possibly limiting the investment risk.

Bhagat and Welch (1995) argue that since R&D investment produces future benefits that are less likely to be realized if the firm becomes distressed, highly levered firms should spend less on R&D. However, they find that the previous year's debt ratio is positively correlated with the current year of R&D activities. They suggest that financial distress costs are not a major determinant for companies. They report a negative relation between current R&D investment and the financial solvency of the firm.

Financial Autonomy translates the solvency of a company. This concept is extremely useful in the assessment of financial structure of the company and on its ability to fulfil their long-term financial commitments.

This concept is extremely useful in the assessment of long-term financial risk in that it provides information on the financial structure of the company and on its ability to fulfil towards their long-term financial commitments. In fact, the greater financial autonomy, the greater part of their applications being financed by equity and, consequently, the lower part being financed with recourse to external or debt financing, i.e. lower the degree of the company's debt. Based on these arguments, we formulated the following hypothesis.

H1. *Corporate financial autonomy increases the probability of firm carry out R&D activities, while high debt inhibits.*

3.1.1.2. Profitability of the company

Profitability as a financial resource and a determinant of financial autonomy, plays a major role in R&D activities. The greater the profits generated by the activity, the greater the accumulation of own capital and greater will be its self-financing capacity, contributing to enhancing the financial autonomy.

Reynard (1979) analysed the correlation between R&D investment and profitability. He stated that the pertinence for the success of R&D investment generally encourages firms to conduct R&D activities, but that excessive R&D expenditures can be a burden to firms because it cuts their profits. Therefore, he argued that the optimal R&D investment be decided from the proportion of net profit to sales. He showed that a decrease of net profit is statistically significant with a decrease of R&D investment.

Myers (1984) stated that firms prefer to finance new investment, first internally with retained earnings, than with debt, and finally issuing new equity calling this the 'pecking order theory'. He considers that this theory assumes a negative correlation between the profitability and the degree of financial leverage.

Coad and Rao (2010) alleged that profitability is an important internal determinant factor for R&D investment. It is expected that profits influence R&D activities, although there is no consensus in this point. Contrary to 'Schumpeterian' perception, Coad and Rao (2010) did not detect any strong association of growth of profits with subsequent investment in R&D once they controlled for sales growth and employment growth. Although they were unable to test the hypothesis that it is the expectation of profits that drives R&D investment.

Lai, *et al.* (2015) only found in Taiwan, one of the three countries (Taiwan, Japan, and Korea) that profitability has a positive and significant correlation with R&D activities. Based on these arguments, we formulated the following hypothesis.

H2. *A higher degree of enterprise revenue or profitability will lead to more active R&D investment activities.*

3.1.2. Physical resources

Physical resources include tangible items that are necessary and available for a business to function. Tangible resources strongly influence R&D investment. Without proper and minimum tangible resources, the enterprise cannot conduct R&D activities.

3.1.2.1. Asset structure

Many scholars identify the relative importance of equipment, manufacturing facilities and buildings firms used in R&D activities. Galende Del Canto and Suárez González (1999) recalls that carrying out R&D activities requires a minimum prior investment in highly sophisticated technical equipment. Highly sophisticated technical equipment also raises the intensity of the capital factor.

When a company has major costs from investments in equipment, it is a capital intensive company. For capital intensive companies, asset structure is represented largely by tangible assets. Based on these arguments, we formulated the following hypothesis.

H3. *A higher depreciation of an enterprise's capital structure will lead to a higher willingness to invest in R&D-related activities.*

3.1.2.2. Company size

There are numerous studies linking company size with R&D activities. Schumpeter (1934) argued for a large-firm advantage in R&D. Economies of scale in R&D or the existence of necessary critical mass make research difficult for small firms. He considers that firm size is an indicator of its power in the marketplace, which would favour innovation by facilitating the appropriation of returns derived from this activity.

Consistent with Schumpeter, Galende Del Canto and Suárez González (1999) suggested that larger companies are more likely to engage in R&D and Fishman and Rob (1999) and Cohen and Levin (1989) also suggested that larger firms invest more in R&D than smaller ones. Since Schumpeter’s work, a vast literature has emerged concerning the effect of the firm size on R&D. Several empirical studies (such as Kamien and Schwartz (1975)) suggest that R&D activity increases more than proportionately with firm size, supporting that size has a positive effect although not linear. Based on these arguments, we formulated the following hypothesis.

H4. *Larger companies increase the probability to invest in R&D activities.*

3.1.3. Intangible resources

Intangible assets, sometimes mentioned to as knowledge assets or intellectual capital, are essentially assets which do not have a "physical or financial embodiment", such as patents, goodwill, software, copyrights, designs, trademarks and brands.

Galende Del Canto and Suárez González (1999) highlights the importance of intangible resources and even considered they are often the most important ones from the strategic point of view, since they are required in the basis of competitive advantage. However, they are also the most difficult resources to observe and evaluate.

3.1.3.1. Goodwill and patent

Goodwill is often referred to as the most intangible of the intangibles because it can only be identified with the business as a whole (Standfield (2005)). Companies find it extremely difficult not only to identify certain types of intangibles but also to assign a value to them in a business combination. As a result, smaller companies do not identify or record intangible assets, consider the process too complex and associate any costs with future benefits too difficult (Marr (2008)). On the contrary, patents, unlike goodwill, can be sold or exchanged individually in the marketplace.

Arora, *et al.* (2008) found that investment activities in R&D and profitable patents have a significant and positive relationship. Varsakelis (2001), in a cross-country study, concluded that countries with strong patent protection invest more in R&D and that the national culture is correlated to R&D investment.

Wang (2010) found there is a close relationship between R&D activities and knowledge generation (with patents as a proxy indicator). Wang (2010) suggests three different broad theories about the purposes of patents "*the first is the invention motivation theory, which posits that the anticipation of patents provides motivation for useful invention. The second is the inducing commercialisation theory, which argues that patents on inventions induce the needed investments to develop and commercialise those inventions. The third can be referred to as information disclosure theory, which states that patents are society's award to individuals or firms who disclose their inventions*" (p. 104). Lai, *et al.* (2015) concluded that an enterprise's goodwill and patents have a close relationship with investments in

R&D. Existing literature considered the same and also that patent rights protection scheme is an impulse to R&D investments. Based on these arguments, we formulated the following hypothesis.

H5. *Better enterprise goodwill and a higher degree of accumulated patents can lead to a higher degree of R&D activities.*

3.1.3.2. Human resources

Human resources and education level were considered significant determinants, and it was found that more developed countries tend to provide stronger patent rights Ginarte and Park (1997) and consequently increased R&D activities.

The talent and strength of employees can lead to successful organizations and the hiring of experienced professionals ensures the mission and goals of the company will be carried out efficiently and with quality.

Galende and Fuente (2003) related the qualified human capital with the cumulative nature of the innovation. They justified that qualified personnel can carry out more intense and continuous research work.

Gustavsson and Poldahl (2003) consider human capital relevant for the innovative process. They supported the model of creative destruction because they found robust evidence for a firm's' human capital to be a significant determinant for R&D spending.

Lai, *et al.* (2015) supported that better human resources can lead to a higher engagement in R&D activities. Based on these arguments, we formulated the following hypothesis.

H6. *Higher qualified human resources increase R&D activities*

3.1.3.3. Business resources

Another intangible resource is the business resources crucial for the reputation and image of the firm. The literature gives important significance to international markets and at this point shows us the importance of competition (Teece (1986)). Competition source in foreign markets is severe instead of domestic market. All firms may benefit from the increased availability of competition, and find that supply of imported technology improves with the increased presence of transnational enterprise. Competition, encourages both the importation and the utilisation of new technology (Braga and Willmore (1991)).

Schumpeter (1934) predicted a negative relation between product market competition and R&D activities. Aghion, *et al.* (1998) showed how various changes in the model set-up may reverse the predicted negative impact of competition on R&D and growth. Based on these arguments, we formulated the following hypothesis.

H7. *Business resources such as an enterprise's export activities will positively affect the investment amount in the enterprise's R&D.*

3.2. Model

We decided to apply binomial logistic regression to predict the relationship between the dependent and the independent variables, following some similar studies (Galende Del Canto and Suárez González (1999) and Lai, et al. (2015)). The main aim is to predict the probability of how our chosen variables influence the level of investment in R&D. Binary logistic regression is mostly used when the dependent variable is dichotomous and the independent variables are either continuous[2] or categorical[3].

Limited dependent variable analysis is appropriate when dependent variable takes binary values ($Y = 1\ or\ 0$). Logit regression models allow us to estimate the Prob. ($Y = 1$) (probability of dependent variable or the probability of carrying out R&D activities) as a function of the type of resources it controls. According to Gujarati (2009), the logit regression model can be written as:

$$P_i = E(Y = 1|X_i) = \frac{1}{1 + e^{-(\alpha + \Sigma_i \beta_i X_i)}} = \frac{1}{1 + e^{-Z_i}} = \frac{e^Z}{1 + e^Z} \qquad \text{Eq. (3.1)}$$

[2] Can be further categorized as either nominal, ordinal or dichotomous.
[3] Also known as discrete or qualitative variables.

where, α is a constant; β_i is the estimated coefficient; X_i is the independent variable; $Z_i = \beta_1 + \beta_2 X_i$. Now it is easy to verify that as Z_i ranges from $-\infty$ to $+\infty$, P_i ranges between 0 and 1 and that P_i is nonlinearly related to Z_i.

According to Park (2013), the binomial logistic model needs to follow these assumptions:

1. Uses maximum likelihood than ordinary least squares estimation. The maximum likelihood estimation is more flexible in the data and analysis because it has fewer restrictions. This model avoids many of the typical assumptions tested in statistical analysis such as:
    - Does not assume normality of variables (both dependent and independent variables);
    - Does not assume linearity between dependent and independent variables;
    - Does not assume measurement level of the independent variable;
    - Does not assume homoscedasticity of the errors;
    - Does not assume normality of the error distribution.
2. The dependent variable should be measured on a dichotomous or binomial scale. Our dependent variable (*RD*) is coded accordingly, being the factor level 1 represented the desired outcome;
3. Can handle non-linear relationship between the dependent and independent variables, because it applies a non-linear log transformation of the linear regression;
4. One or more independent variables, which can be either continuous or categorical;
5. Should have low or no multicollinearity, the independent variables should be independent from each other;
6. Observations and dependent variables should have mutually exclusive and exhaustive categories;
7. The model should be fitted correctly. It should not be over fitted with the meaningless variables included. Also it should not be under fitted with meaningful variable not included;

8. Error terms (residuals) do not need to be multivariate normally distributed although multivariate normality yields a more stable solution. The variance of errors can be heteroscedastic for each level of the independent variables, because logit uses standard distribution errors. Error terms are assumed uncorrelated and need to be independent;
9. Does not require a linear relationship between the dependent and independent variables, it requires that the independent variables are linearly related to the log odds of an event;
10. Requires large sample sizes because maximum likelihood estimates are less powerful than ordinary least squares used for estimating the unknown parameters in a linear regression model.

4. Data and methodology

In this section we will describe our sample and data collection, section 4.1 , the descriptive statistics, section 4.2 and the model used, section 4.3.

4.1. Sample and data collection

Our sample was obtained from Thomson Reuters Datastream in firm level with numerical data collection covering global listed companies figures for 2015. In the following points we will describe our sample and data collection.

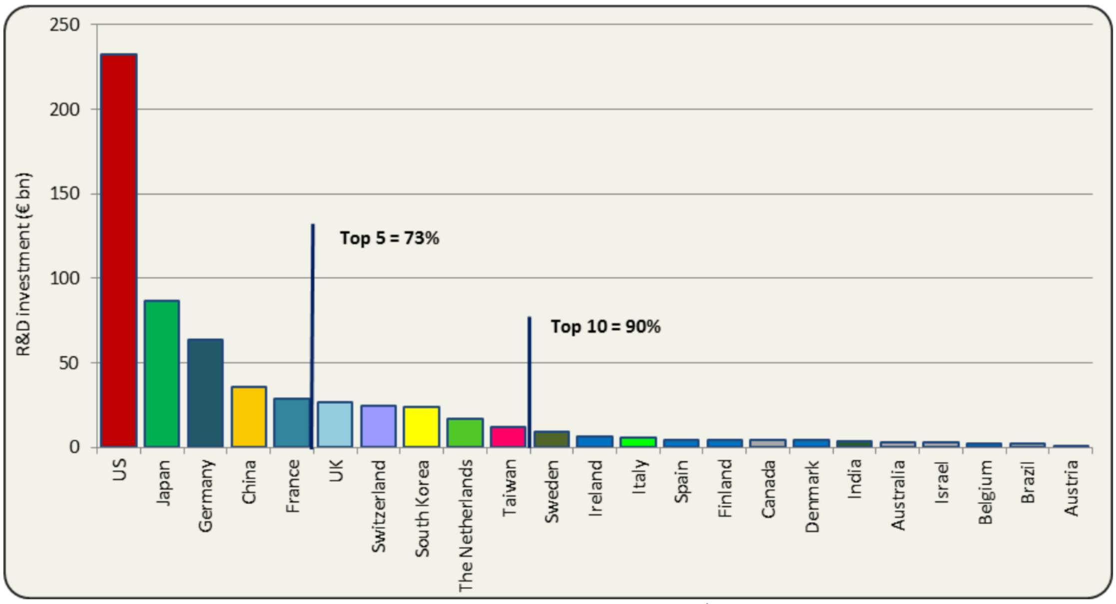
4.1.1. Cross-country study

We decided to run a cross-country regression although researchers are tightly constrained to the assumption of universalism, a tendency to universalize the results (McCartney (2006)). The literature shows us a good example, Galende Del Canto and Suárez González (1999) use a single country, while Lai, *et al.* (2015) assumed parameters identical in a cross-country study. According to the McCartney's view, each individual country in cross-section can be used to elucidate the one underlying universal R&D relation.

Several studies in R&D stated the procedures of OECD (2002) in Frascati Manual that classifies R&D data on a territorial basis in terms of sources of funding and sectors of expenditure as well as in terms of socio-economic objectives, research fields and types of research.

We decided to use Germany and France local firms because according to Guevara, *et al.* (2015) these two countries are on the top European list of R&D investment of 2015 as can be seen in Figure 1.

Figure 1 - R&D investment of the 2015 Scoreboard aggregated by country

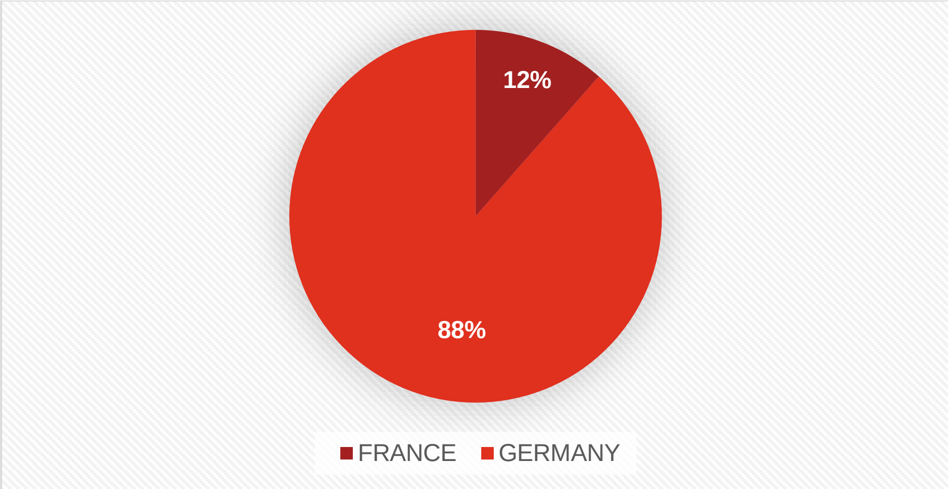


Source: The 2015 EU Industrial R&D Investment Scoreboard, European Commission, JRC/DG RTD from EU R&D Scoreboard 2015.

The two countries which presented the higher number of companies which invested in industrial R&D investment in 2015 were Germany with 136, UK with 135 and France with 86 companies (Guevara, *et al.* (2015)). Unfortunately, we did not have enough data to support a study also for the UK.

Our sample has 64 companies located in France and 492 companies located in Germany.

Figure 2 – Percentage of companies by country in our sample.



Source: Author's own elaboration.

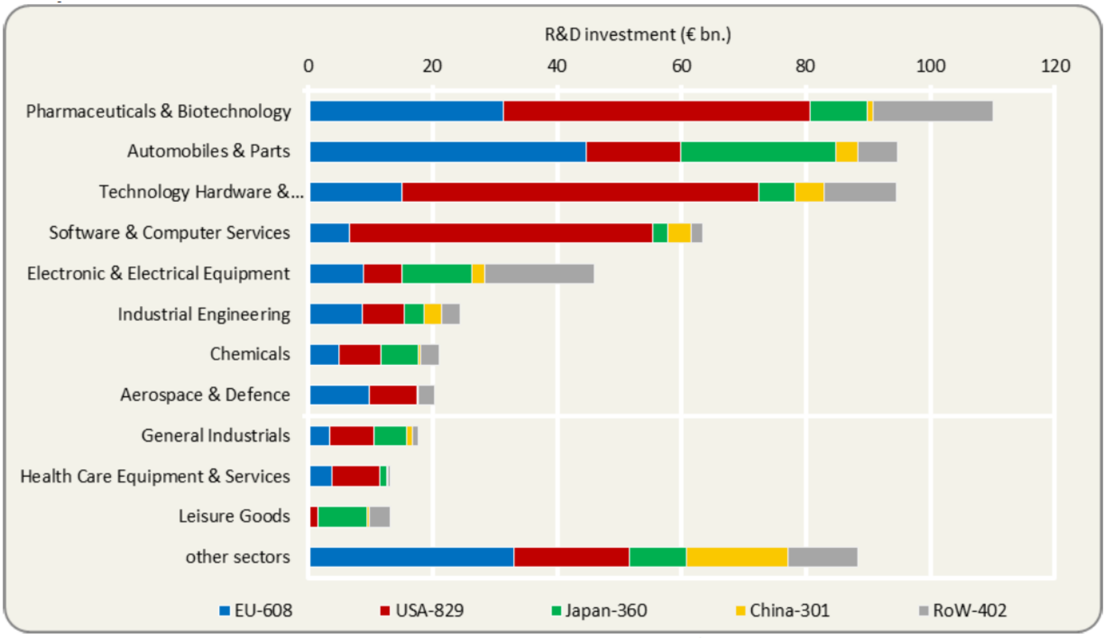
4.1.2. R&D intensity sectors

Most of the studies (for example Lai, *et al.* (2015) and Chen and Miller (2007)) used sectors with R&D intensity, focusing in the manufacturing companies to avoid possible bias. According to Guevara, *et al.* (2015) there is a division between four stages of R&D intensity sectors (R&D as % of net sales)[4]. The first one with high R&D intensity include Pharmaceuticals & biotechnology; Health care equipment & services; Technology hardware & equipment; Software & computer services, Aerospace & defence and Leisure Goods. The medium-high R&D intensity included Electronics & electrical equipment; Automobiles & parts; Industrial engineering; Chemicals; Personal goods; Household goods; General industrials; Support services sectors. The medium-low R&D intensity sectors include mainly Food producers; Beverages; Travel & leisure; Media; Oil equipment; Electricity; Fixed line telecommunications. Lastly, the low R&D intensity sectors include mainly Oil & gas producers; Industrial metals; Construction & materials; Food & drug retailers; Transportation; Mining; Tobacco; Multi-utilities.

Guevara, *et al.* (2015) also presented, in Figure 3, the distribution of R&D investment by industrial sector and region.

[4] This classification takes into account the R&D intensity of all companies aggregated by ICB 3-digit sectors: High above 5%; Medium-high between 2% and 5%; Medium-low between 1% and 2% and Low below 1%. Some sectors are adjusted to compensate for insufficient representativeness of the Scoreboard using the OECD definition of technology intensity for manufacturing sectors (ISIC REV 3. Technology intensity definition, OECD, 7 July, 2011.)

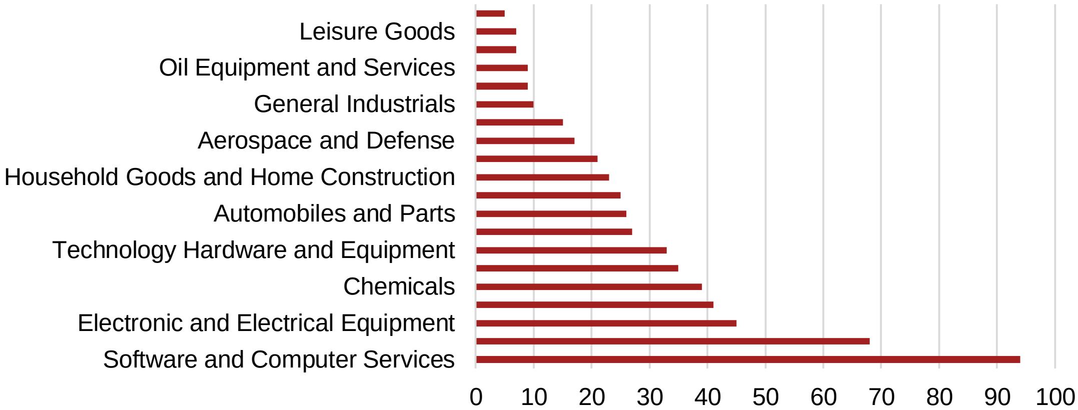
Figure 3 - R&D ranking of industrial sectors and regional shares for the world's top 2500



Source: The 2015 EU Industrial R&D Investment Scoreboard, European Commission, JRC/DG RTD.

In order to obtain a homogeneous accounting sample, we decided to exclude banks, insurance companies and financial services. Therefore we focus our sample in the following sectors: Software & Computer Services, Electronic & Electrical Equipment, Chemicals, Technology Hardware & Equipment, Automobiles & Parts, Household Goods and Home Construction, Aerospace & Defence, General Industrials, Oil Equipment and Services and Leisure Goods. We only excluded from our sample the low and medium-low R&D intensity sectors (according to Guevara, *et al.* (2015)). Our sample's sectors are expressed in the Figure 4.

Figure 4 - Number of companies by sectors in our sample.



Source: Author's own elaboration.

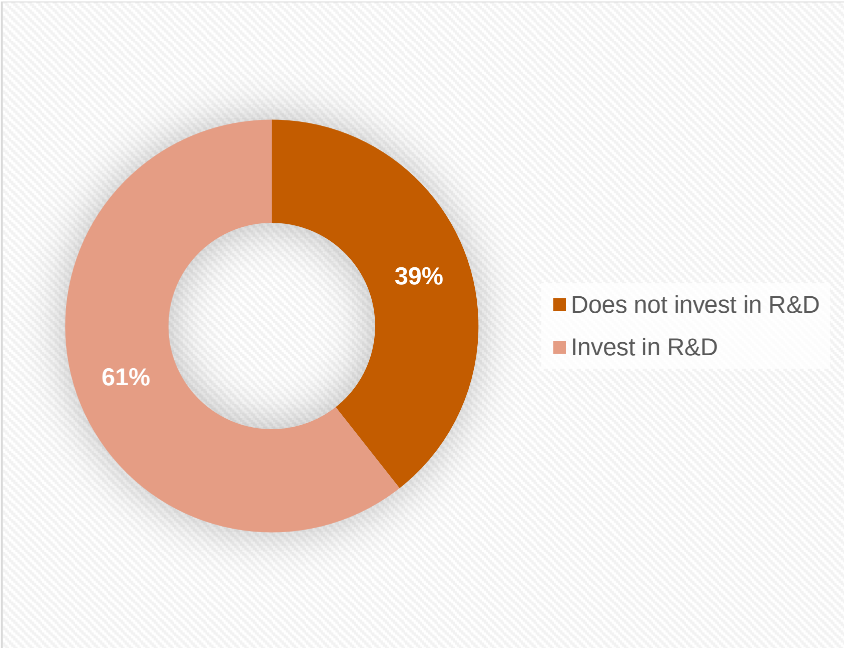
Lack of data availability reduced the number of firms. After some analysis, we realized that our sample held some errors[5], so we excluded or adjusted those data. After these interventions on data, the obtained sample comprised of 556 firms.

4.1.3. Operation measures

In section 3.1 we describe our research hypothesis, in which we detail all the determinants used according to our previous approach and previous literature. The dependent variable in the analysis is firms' R&D binary dummy (1 = if firm reports any R&D activity in 2015; 0 = otherwise). Our sample comprises 61% of companies which do not report any R&D activity and 39% of companies that reports R&D activity, as shown in Figure 5.

[5] Our ratio HR in some companies presented a two digits results. At some point we realized that the data obtained by Datastream was not in a million units and we had to adjust this variable.

Figure 5 - Percentage of companies on our sample with and without R&D investment



Source: Author's own elaboration

The following points have the same connection to our hypothesis.

4.1.3.1. Financial resources

In this section we will introduce three ratios. Two of the most popular calculations, the equity ratio and the leverage ratio, allow us to analyse the influence of the firm's capital structure as determinant of the carrying out of R&D activities. The third one allows us to analyse profitability.

The equity ratio is an investment solvency ratio that measures the amount of assets financed by owners' investments and highlights the concepts of a solvent and sustainable business. Companies with higher equity ratios show new investors and creditors that believe in the company and are willing to finance it with their investments. The ratio is expressed and calculated as follows:

$$Equity\ (FIN) = \frac{Total\ Equity}{Total\ Assets} \qquad \text{Eq. (4.2)}$$

Higher ratios are typically favourable to perform R&D activities so we expect a positive sign.

The leverage ratio is a debt ratio that measures how in debt is a company:

$$Leverage\ (LEV) = \frac{Long\ and\ short\ term\ liabities}{Equity} \qquad \text{Eq. (4.3)}$$

Higher positive ratios are typically indicators that would be harmful to perform R&D activities so we expect a negative sign. These two ratios are directly related to hypothesis 1.

In order to measure profitability, the return on assets ratio (ROA) was used. This ratio shows how efficient management is at using its assets to generate earnings. The formula for return on assets is:

$$ROA\ (PRO) = \frac{Net\ Income}{Total\ Assets} \qquad \text{Eq. (4.4)}$$

Higher ratios mean that companies are better at converting its investment into profit, therefore we expect a positive sign, explaining the formulated hypothesis 2.

4.1.3.2. Physical resources

The next two ratios allow us to analyse the physical resources of the company. The following ratio measures the capital intensity:

$$Amortisation\ (PHR) = \frac{Depreciation\ Costs}{Net\ Turnover} \qquad \text{Eq. (4.5)}$$

We chose the logarithm of assets to measure the size of the company.

$$Size\ (SZ) = \log\ (Assets) \qquad \text{Eq. (4.6)}$$

The formulas above correspond to our hypothesis number 3 and 4. We expect a positive sign in both variables.

4.1.3.3. Intangible resources

Intangible resources have a high importance in our study. However, it is not easy to quantify it. Perhaps it would be easier to measure goodwill and patent because companies have a close relationship with investments in R&D. We use the ratio as follows:

$$Goodwill\ and\ Patents\ (IR) = \frac{Goodwill\ and\ Patents}{Total\ Intangible\ Other\ Assets} \qquad \text{Eq. (4.7)}$$

To reach the qualification of human resources our proxy is the average wage, which can be seen below, although a very restrictive formula.

$$Average\ Wage\ (HR)\ in\ euros = \frac{Labour\ Costs}{Total\ Employees} \qquad \text{Eq. (4.8)}$$

An international company has higher potentials to grow, so we measure through binomial values if the company exports, as can be see below.

$$Exports\ (EXPT) = dichotomic\ variable = \begin{cases} 1 = firm\ exports \\ 0 = \ otherwise \end{cases} \qquad \text{Eq. (4.9)}$$

A positive effect is also expected for goodwill and patents, average wage and exports. These variables are respectively linked to hypothesis 5, 6 and 7.

Table 1 summarises the research hypothesis formulated, the respective variables and the expected signal towards R&D activity.

Table 1 – Proposition and hypothesis

	Hypothesis	Variable Code	Signal
H1	Corporate financial autonomy increases the probability of firm carry out R&D activities, while high debt inhibits.	*FIN*	POSITIVE
		LEV	NEGATIVE
H2	A higher degree of enterprise revenue or profitability will lead to more active R&D investment activities.	*PRO*	POSITIVE
H3	A higher depreciation of an enterprise's capital structure will lead to a higher willingness to invest in R&D-related activities.	*PhR*	POSITIVE
H4	Larger companies increase the probability to invest in R&D activities.	*SZ*	POSITIVE
H5	Better enterprise goodwill and a higher degree of accumulated patents can lead to a higher degree of R&D activities.	*IR*	POSITIVE
H6	Higher qualified human resources increase R&D activities	*HR*	POSITIVE
H7	Business resources such as an enterprise's export activities will positively affect the investment amount in the enterprise's R&D.	*EXPT*	POSITIVE

Source: Author's own elaboration

4.2. Descriptive statistics

Table 2 exhibits the mean, median and standard deviation for each variable, allowing to analyse the descriptive statistics for each variable.

Table 2 - Descriptive statistics of the sample

	FIN	LEV	PRO	PhR	SZ	IR	HR
Observations	556	556	556	556	556	556	556
Mean	0,38	2,01	4,12	0,13	6,22	0,43	82.790
Median	0,35	1,32	5,24	0,09	6,17	0,39	66.255
Std. Deviation	0,21	2,54	16,56	0,19	1,07	0,30	71.579

Note: Values with more than two decimal digits have been rounded up to two decimal places instead of *HR* column, rows mean, median and Std. Deviation, which was rounded to the unit level.

Source: Author's own elaboration

Now we will look into each single variable, analyse the most relevant values and compare to our literature (with special focus on Galende Del Canto and Suárez González (1999) and Galende and Fuente (2003)). Our variable, *FIN* displays values in accordance with our literature, showing evidences that on average our companies' sample has a positive level of financial autonomy. The leverage ratio (*LEV*) suggests that on average, our companies' sample has a significant level of debt, although our median is lower but still with a significant level of debt and with standard deviation (above the mean) indicating a great level of dispersion. Profitability (*PRO*) on average shows a high ratio however it has a gigantic level of standard deviation, much higher than mean. Physical Resources (*PhR* and *SZ*) have congruent values, although very different from Galende Del Canto and Suárez González (1999). *IR* exhibits a mean of 0,43, which tells us that in mean our companies' sample has invested in goodwill and patents. *HR* exhibits an expected result taking into account that our study is based on Germany and France and our sample is R&D intensity.

The *EXPT* variable is not present in Table 2 for it is a dichotomous variable, without relevance in this analysis.

4.3. Model

In line with some previous studies in the literature, using similar data and what was already said in section 3.2., we will estimate the following model:

$$P(RD_i = 1) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \times IR_{iy} + \beta_2 \times EXPT_{iy} + \beta_3 \times FIN_{iy} + \beta_4 \times LEV_{iy} + \beta_5 \times PRO_{iy} + \beta_6 \times PHR_{iy} + \beta_7 \times HR_{iy} + \beta_8 \times SZ_{iy})}} \qquad \text{Eq. (4.10)}$$

Where *RD* is the dependent variable, being *P(RD)* the probability of the dependent variable; *IR*, *EXPT*, *FIN*, *LEV*, *PRO*, *PhR*, *HR* and *SZ* are the independent variables, chosen and mentioned in section 4.1.3; $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$ are the regression coefficients of independent variables; *i* is the correspondent company; *y* is the year of 2015.

5. Empirical results

In this section we will address the results of the collected data and the interpretation and discussion of results achieved in empirical research on firm's decision to carry out R&D activities. It was used for main statistical analysis IBM Statistics SPSS (Version 21.0) and with support of Eviews (version 8) in certain analysis (such as McFadden R^2).

The significance level (α) was set at 0,05 (5%) for all hypothesis tests. The decision criterion used is in accordance with the requirement of the current scientific community: when the p-value found is less than or equal to α, reject the H_0 (null hypothesis), considering that the difference found is significant. If found p-value is greater than α, we retain the null hypothesis (H_0) and decided there was no significant difference (alternative hypothesis rejection, H_1).

First, we will address the multicollinearity problem, that occurs when two or more independent variables in the model are approximately determined by a linear combination of other independent variables in the model. As already mentioned, the condition of non-multicollinearity is very important for this model, however there is not a single method of detecting it or measuring its strength Gujarati (2009).

Table 3 presents the correlation coefficients between variables. Coefficient of correlation (r) is a measure for statistical correlation, and may range from $-1 \leq r \leq +1$. This measure gives the strength of a linear relationship between variables. In general terms, for a positive association the coefficient must be positive ($r > 0$), otherwise for a negative association the coefficient must be negative ($r < 0$) while for no relationship or the variables are independent and not related, the coefficient must be zero ($r = 0$). When $r = +1$ describes a perfect positive correlation and $r = -1$ describes a negative correlation, the sign of the correlation provides direction only.

Taking into consideration the suggestion of Evans, *et al.* (1996) we can conclude that in our sample (Table 3) most of Pearson's correlations are very weak. Correlation between *SZ* and *FIN* (-0,314), *SZ* and *LEV* (0,232) and even *PhR* and *PRO* (-0,354) have weak strength. There is a moderate strength between two variables *LEV* and *FIN* (-0,558). As showed below the corresponding p-values are p < 0,001 thus we can say that they are all significant.

Table 3 – Pearson's correlation matrix

	FIN	LEV	PRO	PHR	SZ	IR	HR	EXPT
FIN	1	-0,558**	0,040	0,013	-0,314**	0,042	-0,102*	-0,007
LEV	-0,558**	1	-0,103*	0,052	0,232**	0,063	0,009	0,013
PRO	0,040	-0,103*	1	-0,354**	0,185**	-,086*	0,000	0,080
PHR	0,013	0,052	-0,354**	1	-0,095*	-0,047	-0,003	-0,122**
SZ	-0,314**	0,232**	0,185**	-0,095*	1	-0,040	0,166**	0,143**
IR	0,042	0,063	-0,086*	-0,047	-0,040	1	-0,107*	0,100*
HR	-0,102*	0,009	0,000	-0,003	0,166**	-0,107*	1	0,014
EXPT	-0,007	0,013	0,080	-0,122**	0,143**	0,100*	0,014	1

Note: ** Correlation is significant at the 0,01 level (2-tailed); * Correlation is significant at the 0,05 level (2-tailed). Values with more than three decimal digits have been rounded up to three decimal places.

Source: Author's own elaboration

The results of the logistic regression model can be now presented through Table 4.

Table 4 – Logistic model

Variable	Coefficient	Std. Error	z-Statistic	Prob.	Exp. (Coeff.)
C	-3,8139	0,7727	-4,9360	0,0000	0,0221
FIN	2,9973	0,6841	4,3817	0,0000	20,0310
LEV	0,0283	0,0470	0,6012	0,5477	1,0287
PRO	-0,0173	0,0082	-2,1133	0,0346	0,9829
PhR	-1,114	0,7911	-1,4080	0,1591	0,3283
SZ	0,4523	0,1020	4,4323	0,0000	1,5718
IR	-0,78	0,3288	-2,3719	0,0177	0,4584
HR	-0,0001	0,0000	-2,3499	0,0188	1,0000
EXPT	1,3114	0,2594	5,0549	0,0000	3,7111
McFadden R^2			*0,1014*		
LR statistic			*75,6199*		
Prob (LR statistic)			*0,0000*		
Obs with RD = 0			*219*		
Obs with RD = 1			*337*		
Total obs			*556*		

Note: Values with more than four decimal digits have been rounded up.

Source: Author's own elaboration.

At first sight, we can see that our study presented a significant explanatory power with the p-value so closed to zero, which will be analysed further on.

The first column contains the estimated coefficients on each independent variable with the correspondent probability as known as p-value (p). In the second column we can see the standard error and then in the third column the Z statistic. The fourth column shows the p-value estimation and in the last one the exponential coefficient (or odds). Other statistics related to the logistic regression are shown below.

In the first column, a positive coefficient for an independent variable relates to increase of probability that a firm will carry out R&D activities. Negative coefficients relate to a decrease of probability that a firm will carry out R&D activities. Before proceeding with this analysis, we must certify that all our variables are significant, so we must look at the fourth column ("Prob.").

There are six statistically significant variables (p-value < 0,05) *FIN*, *PRO*, *SZ*, *IR*, *HR* and *EXPT*, three of them (*FIN*, *SZ* and *EXPT*) having p-value < 0,01, unlike *LEV* and *PhR* which are statistically insignificant.

The *IR* coefficient of -0,7800 means that for one unit change in *IR*, the log odds of R&D activities decreases by 0,78, i.e. *IR* for every one unit change in *IR* the log odds of R&D activities decreases 0,78 units, ceteris paribus, suggesting a negative relationship between the two. While the indicator variable *EXPT* has a slightly different interpretation, if a company exports versus a company that does not export, the log odds of R&D activity increases 1,3114. As you can see, *PRO*, *IR* as *HR* have negative effect. However, *FIN*, *SZ* and *EXPT* have positive effect on the logit.

A more meaningful interpretation is in terms of odds, which are obtained by taking the antilog of the distinct slop coefficients, which can be seen in the last column (Exp. (Coeff.) known as odds ratios of the individual coefficient). As you can see the following equation allows us to reach the odds:

$$\text{ODDS} = e^{a+bX} \qquad \text{Eq. (5.11)}$$

Companies that exports (*EXPT* = 1) are more than 3,71 ($\approx e^{1,3114}$)[6] times likely to carry out R&D activities than companies that do not export, ceteris paribus. The most significant is *FIN*, companies with higher financial autonomy have more than 20 times likely to carry out R&D activities, ceteris paribus. And we can easily convert odds into probabilities (%), throughout the following formula and display on Table 5.

$$\hat{Y} = \frac{ODDS}{1 + ODDS} \times 100 = \% \qquad \text{Eq. (5.12)}$$

As mentioned above, the errors (i.e., residuals) from the linear probability model violate the homoscedasticity and normality of errors assumptions of ordinary least squares regression, resulting in invalid standard errors and hypothesis tests (Long and Freese (2006)).

In the third column we can analyse the value of Z-Statistic. Z-value can be too big in magnitude (i.e., either too positive or too negative), it indicates that the corresponding true regression coefficient is not 0 and the corresponding X-variable matters. A good rule of thumb is to use a cut-off value of 2 which approximately corresponds to a two-sided hypothesis test with a significance level of $\alpha = 0{,}05$ (5%). So, for the *LEV* variable, the z-value is 0,6012 which is not large enough to provide strong evidence that *LEV* matters validating the p-value already analysed.

Now we will test how well the model describes the response variable with the goodness of fit. Assessing goodness of fit involves investigating how close values predicted by the model are to the observed values.

To test the null hypothesis that all the slope coefficients are simultaneously equal to zero, we pay attention to the likelihood ratio (LR) statistic. In our sample LR is 75,62, whose p-value is approximately 0[7], which is very small and expressed together, all the variables have a significant impact on R&D activities. The LR test is based on -2 log likelihood ratio that will be analysed ahead. LR is a test of the significance of the difference between the LR for the model with all the predictors (called model chi square) minus the LR for

[6] The value obtained was 3,71106909175.
[7] The value obtained was 0,0000000000003705924456198773.

baseline model with only a constant in it. This ratio measures the improvement in fit that explanatory variables make compared to the null mode.

Odds ratio, confidence interval of odds and Wald test are displayed in Table 5. The Wald test is similar to the LR test but in here we tested the hypothesis that each β = 0. These two types of chi-square tests are asymptotically equivalent. If the Wald statistic is significant (i.e., less than 0,05) then the parameter is significant in the model. The square of this Z statistic is approximately a X^2 statistic with one degree of freedom. So we can say that FIN, SZ and EXPT have the larger effect in the model as expected, following the results of p-value.

The 95% confidence interval (C.I.) for the odds lead us to think that there is a statistically significance associated between the variables *FIN*, *PRO*, *SZ*, *IR*, *HR*, *EXPT* and the R&D activities. C.I. does not include the value 1, so we reject H_0. The larger model is better *FIN*, *PRO*, *SZ*, *IR*, *HR* and *EXPT* are still good predictors.

Table 5 – Odds ratio, confidence interval and Wald test.

Variable	Odds Ratio	95% C.I. for Exp. (Coeff.)		Wald test Chi-square
		Lower	Upper	
C				24,3643
FIN	0,9525	5,241290	76,554074	19,1990
LEV	0,5071	0,938155	1,127876	0,3615
PRO	0,4957	0,967291	0,998749	4,4661
PhR	0,2471	0,069637	1,547493	1,9826
SZ	0,5000	1,286920	1,919755	19,6455
IR	0,3143	0,240634	0,873325	5,6257
HR	0,5000	0,999994	0,999999	5,5219
EXPT	0,7877	2,231951	6,170405	25,5521

Note: Values with more than four decimal digits from columns "Odds Ratio" and "Wald test Chi-square" and with more than six decimal digits from columns "95% C.I. for Exp. (Coeff.)" have been rounded up. All variables referring to column "Wald test Chi-square" has a degree of freedom of 1.

Source: Author's own elaboration.

In the logit model, the conventional measure of goodness of fit, R^2, is not particularly meaningful, so instead we use pseudo R^2. In accordance to Allison (2014), who disregards the importance Cox-Snell R^2 in binomial mode (although analysed ahead), we decided to analyse the McFadden R^2 with a value of 0,1014. Like R^2, McFadden R^2 also ranges between 0 and 1, with values closer to zero indicating that the model has no predictive power, which lead us to conclude than this model has a predictive power.

Another comparatively simple measure of goodness of fit is the count R^2 which is defined as the number of correct predictions divided by the total number of observations. In our sample the count R^2 is 67%[8], visible in Table 6. We could reach this value through our Appendix 2, in which out of 556 observations, there were 183 (= 556 - 373) incorrect predictions. By linking these results with those presented by the discriminant analysis, we conclude that the logit regression is more robust in the correct classification of cases.

Table 6 – Expectation-Prediction Evaluation for Binary Specification

	Estimated Equation			Constant Probability		
	RD = 0	RD = 1	Total	RD = 0	RD = 1	Total
P (RD = 1) ≤ C	78	42	120	0	0	0
P (RD = 1) > C	141	295	436	219	337	556
Total	219	337	556	219	337	556
Correct	78	295	373	0	337	337
% Correct	35,62	87,54	67,09	0	100	60,61
% Incorrect	64,38	12,46	32,91	100	0	39,39
Total Gain*	35,62	-12,46	6,47			
Percent Gain**	35,62	NA	16,44			

Note: Success cutoff: C = 0,5

*Change in "% Correct" from default (constant probability) specification

**Percent of incorrect (default) prediction corrected by equation

Source: Author's own elaboration.

The model thus provides meaningful information for identifying important determinants of entry mode choice. The table below, (Table 7) summarizes the tests performed according to the model.

[8] The value obtained was 0,670863309352518 and is a result of $= \frac{556-183}{556}$.

Table 7 – Goodness-of-fit tests

Hosmer and Lemeshow Test		-2 Log likelihood[c]	Cox & Snell R^2	Nagelkerke R^2	Omnibus Tests	
X^2	Prob.				X^2	Prob.
4,624	0,797	669,925[b]	0,127	0,172	75,620	0,000

Note: a. degrees of freedom equal to 8; b. Estimation terminated at iteration number 5 because parameter estimates changed by less than 0,001; c. Initial -2 Log Likelihood: 745.545

Source: Author's own elaboration.

The Hosmer–Lemeshow test which is commonly used to assess the goodness-of-fit, can be tested for any number of explanatory variables, which may be continuous or categorical. The implicit idea is to relate the test adjusted expected values (fitted) to actual values by observations of groups, divided into approximately equal size subgroups. The model can be rejected when the differences are great, they provide a tiny adjustment (fit) to the data. Briefly, the Hosmer-Lemeshow with X^2 distribution is used to test group observations, based on the expected probability that RD = 1. This statistical test has an approximate X^2 distribution with 8 degrees of freedom and a p-value = 0,797 ($p > 0,05$) which indicates that the overall numbers of R&D activity are not significantly different from those predicted by the model. To be a good fitted model, the value of X^2 should be not significant. According to Table **7**, we can conclude this requirement is verified and that, the overall model fit is good.

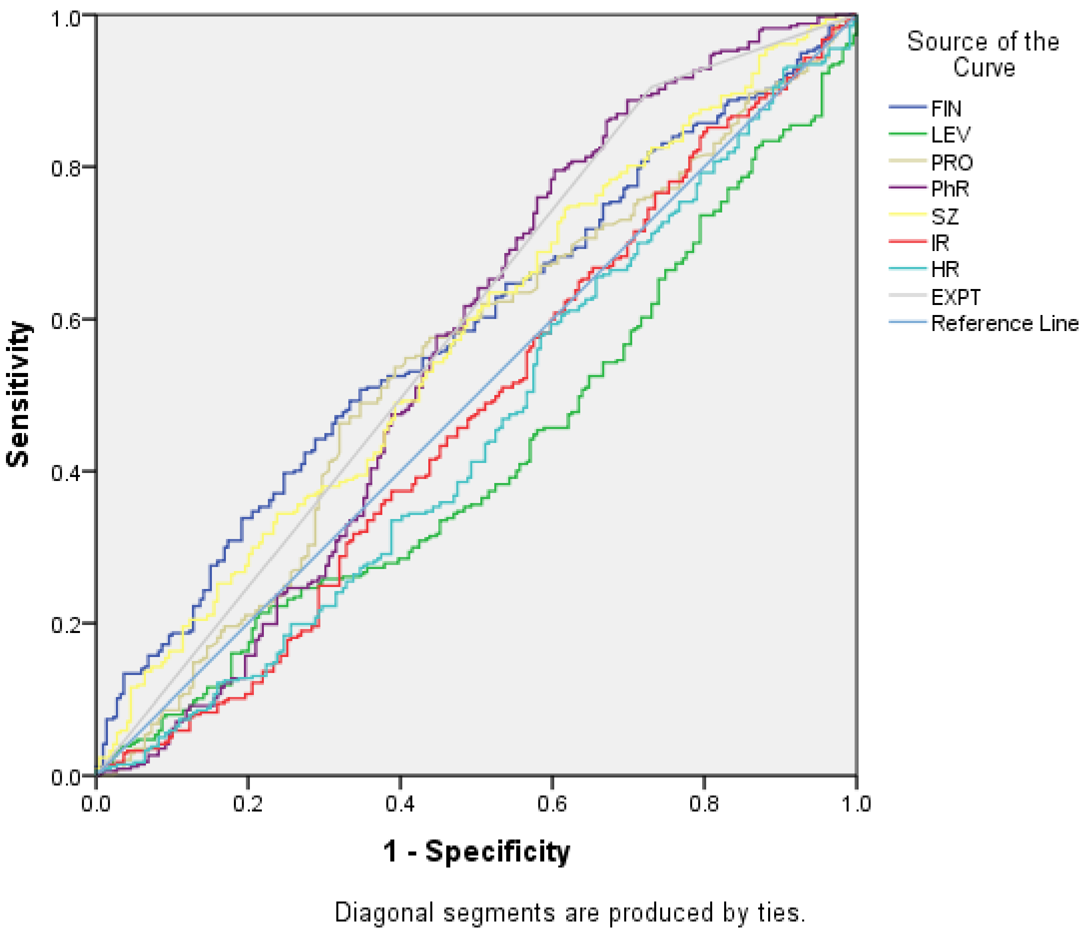
The Omnibus tests of model coefficients (as previously seen as LR – likelihood ratio) is used to check whether the explained variance in a set of data is a statistical improvement over the unexplained variance. It uses X^2 tests to see if there is a significant difference between the Log-likelihoods (specifically the -2 Log likelihood) of the baseline model and the new model. If the new model has a significantly reduced -2 Log likelihood compared to the baseline then it suggests that the new model is explaining more of the variance in the outcome and is an improvement. In our case this fact happens with a -2 Log likelihood baseline model of 745 and a -2 Log likelihood new model (with all explanatory variables) of 669.

In Omnibus tests, the X^2 is highly significant (X^2 = 75,620, df = 8, p < 0,000) which leads to rejection of the null hypothesis. We can conclude that our new model is significantly better, the p < 0,001 indicates the accuracy of the model improves when we add our explanatory variables.

With respect to the test and Cox Snell, the higher the value, the better the fit of the model. Since this test cannot achieve the maximum of one, we often resorts to the Nagelkerke test. The Nagelkerke's R^2 (circled) suggests the model explains roughly 17% of the variation in the outcome. Notice how the two versions (Cox & Snell and Nagelkerke) do vary, so we should interpret as between 13% and 17% of the variation in R&D activities can be explained by the model This shows that these R^2 values are approximations and should not be overly emphasized.

Although the literature queried did not mention the ROC (Receiver Operating Characteristic) curve, we decided to compute such curve to find the sensitivity (the ability of the model to predict an event correctly) vs specificity derived from several cut-points for the predicted value.

Figure 6 – ROC curve



Note: The test result variables: *PRO*, *IR*, *EXPT* has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

Source: Author's own elaboration.

The position of the ROC on the graph reflects the accuracy of the diagnostic test. It covers all possible thresholds (cut-off points). The ROC of a perfect diagnostic technique is a point at the upper left corner of the graph, where the TP proportion is 1,0 and the FP proportion is 0. The Area Under the Curve (AUC), also referred to as index of accuracy (A), or concordance index, c, in SAS, is an accepted traditional performance metric for a ROC curve. The higher the area under the curve the better prediction power the model has. Our model displays that variables *LEV*, *IR* and *HR* have AUC under 0,5 which does not show any reliable prediction. However, *FIN*, *PRO*, *PhR*, *SZ* and *EXPT* variables have AUC above 0,5 which show some kind of prediction in the model.

Now that we know this model has an overall good fit, we can compare our expectations with the actual values of each variable and the correspondent hypothesis. The significant indicators of individual independent variables can partially predict the impact of financial autonomy (*FIN*) on R&D activities. The coefficient is positive and significant as we expected and with an index of accuracy over 0,58. Instead, leverage (*LEV*) has an insignificant coefficient, supported by high level of correlation with financial autonomy. These two variables support partially our hypothesis 1 (H_1), that corporate financial autonomy increases the probability of firms' R&D activities.

Regarding financial profitability (*PRO*), we did expect a positive sign of this coefficient although we do have a negative coefficient that is significant. Financial profitability (*PRO*) reveals an index of accuracy of 0,54. As Lai, *et al.* (2015), which also had a negative value although not significant, we can conclude that this variable could be influenced by physical resources which have a correlation, albeit a weak one. This variable is reversely supported (H_2).

Concerning physical resources (*PhR*), the coefficient model is not significant rejecting hypothesis 3 (H_3). For size (*SZ*) and business resources (*EXPT*) variables, the coefficients are positive and significant ($P < 0,01$) as expected, supporting hypotheses 4 (H_4) and 7 (H_7) respectively. For size (*SZ*) and business resources (*EXPT*), the index of accuracy is respectively 0,57 and 0,59.

Regarding intangible resources (*IR*) and human resources (*HR*) we do have a negative coefficient, contrary to expectations but significant coefficients. Unfortunately, the index of accuracy of intangible resources (*IR*) and human resources (*HR*) does not reach 0,5, meaning it failed to predict R&D activities. These two variables reversely support hypotheses 5 (H_5) and 6 (H_6) respectively.

The following section describes all the conclusions to this study and the limitations of our empirical study. Lastly, possible paths of further investigation in this field will be presented.

6. Conclusion

In this section we will address the main results achieved in this empirical investigation, some study limitations and for last part of this section dedicated to possible paths of further investigation in this field.

Nowadays, innovation has become a source of competitive advantage to face the competitive environment. In order to better guide entrepreneurs we analysed important financial ratios as determinants of firm's R&D activity. Financial ratios were defined, based on literature review, which revealed to be effective, achieving accuracy levels of 67% in the model.

This empirical study explores how leverage (*LEV*), profitability (*PRO*), physical resources (*PhR*), size (*SZ*), intangible resources (*IR*), human resources (*HR*) and business resources (*EXPT*) are potential determinants of firm's R&D activity. We focused on the determination of financial characteristic resources of firms carrying out R&D activities, in contrast with those that do not.

The model was applied to a sample of 556 firms located in France and Germany from 2015. In our study we show evidence concerning the extent to which firms' innovative activities are explained by their internal resources and factors. The general proposition seems to be reasonable and the internal factors affect the configuration of the firm's innovative process.

The logistic model allowed us to estimate the probability of firms carrying out R&D activities. The results indicate that six variables in this study were statistically significant; financial autonomy (*FIN*), profitability (*PRO*), size (*SZ*), intangible resources (*IR*), human resources (*HR*) and business resources (*EXPT*) affecting the odds of firms carrying out R&D activities. Two of our hypotheses were fully supported, one was partially supported, three reversely supported and two rejected. These results tell us that extending our sample to all sectors that invest heavily in R&D may change some perspectives of previous studies which focus on manufacturing companies.

Our study suggest that size and R&D activity have a positive relationship, consistent with Schumpeterian view. Among the seven hypotheses, firms' entrepreneurs should focus in financial resources and internationalization as being the most important internal resources that affect the probability of implementing internal investment activities in R&D. Like Lai, *et al.* (2015), Galende Del Canto and Suárez González (1999) and Galende and Fuente (2003) this study also suggest the importance of strategic decisions behind R&D investments.

Negative signs associated with coefficients of human resources and profitability variables was already detected by Lai, *et al.* (2015). In our study, we may conjecture that this could be related with the industries included in the sample. Companies from industries such as software development typically avoid investing heavily in such assets and are not regarded as capital intensive, despite having R&D activities.

This empirical study has some limitations and we could start with how R&D is measured. All studies about innovation have, most of the time, some kind of issues measuring innovation. Our study tries to override this limitation by choosing a binomial value which allow us to measure the probability to carry out R&D activities. However, reducing a metric variable to dichotomous level misses a lot of information.

Using a logistic model, we had to ensure a relatively large sample for the accuracy of results which deems, the interpretation of the estimates coefficients not so immediate.

Unfortunately, our database limited our research and it was not possible to guaranty a considerable sample for the amount spent on R&D for each company. The same happened when analysing companies in the UK, which initially was part of our study but had to be removed due to the lower number of companies in the sample. In relation to the quality, the human resources (*HR*) variable showed some disproportionate values, which was detected to be data error but later rectified.

For future research it is recommended to analyse more countries simultaneously with similar culture and rules about R&D using quantitative variables and without focusing on manufacturing companies because it was already done. There is a lack of studies that includes not only the manufactured companies but also service companies (which may have a technological focus). Other variable that should be included in this type of study (but, in our study, it was not for lack of data) is the public support or public subsidy which could have a great influence. This should be a subject for further investigation in the future.

7. Appendices

Appendix 1 – Main studies

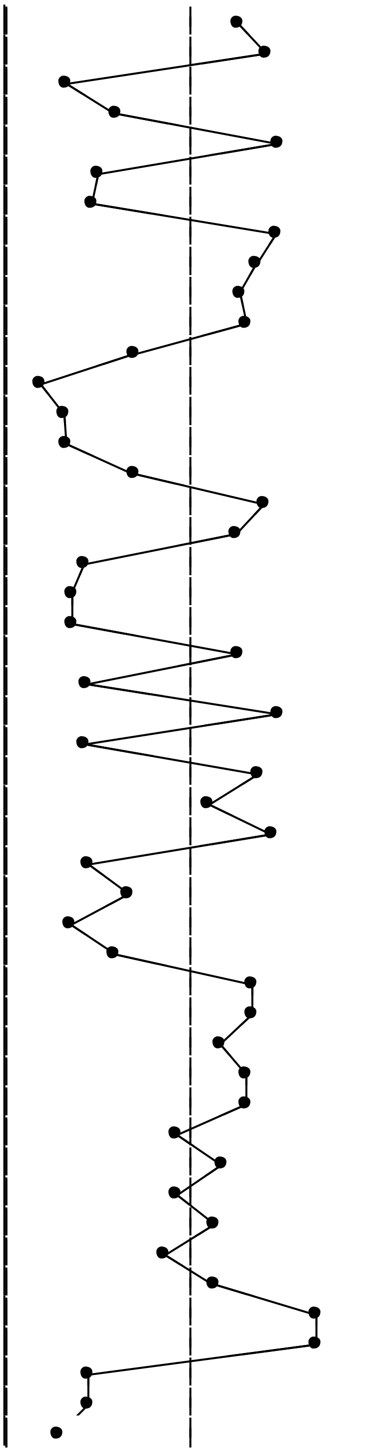
Determinants of R&D Investment	Model / Statistical Analysis	Country of Study	Industrial Sector	Authors
Financial autonomy, leverage, assets, amortisation, average wage and exports.	Logistic Regression Model	Spanish region of Castilla and León	With a turnover of over 480 million pesetas.	Galende Del Canto, Jesús; Suárez González, Isabel (1999)
Knowledge/innovation, public support, demand pull, source of information, appropriable conditions (protection), international competition and size (employees).	Tobit Regression Model	France, Germany, Spain, and the UK	Textile, wood, paper, chemicals, plastic and rubber, non-metallic, basis metals, machinery, electrical, vehicles and non-classified industries.	Griffith, Rachel; Huergo, Elena; Mairesse, Jacques; Peters, Bettina (2006)
Size (employees), age, public funds, foreign capital, cooperation, product innovation, process innovation, environmental impact, norms and regulations.	Tobit Regression and Probit Regression	Spain	Energy Industry	Costa-Campi, M. T.; Duch-Brown, N.; Garcia-Quevedo, J. (2014)
Size (in sales), export share, total national share (in total sales), foreign ownership share, part of bigger firm, public subsidy from EU, costumer and SME.	Binary dependent variable in a generalized linear mixed model	Turkey	Multiple Industries (exclude Agricultural Sector)	Limanlı, Ömer (2015)
Financial Autonomy, profitability, size, capital structure, intangible resources, human resources and business resources.	Logistic Regression Model	Taiwan, Japan and Korea	Manufacturing companies (SIC code 20–39)	Lai, Yung-Lung; Lin, Feng-Jyh; Lin, Yi-Hsin (2015)

Source: Author's own elaboration.

Appendix 2 – Residual plot

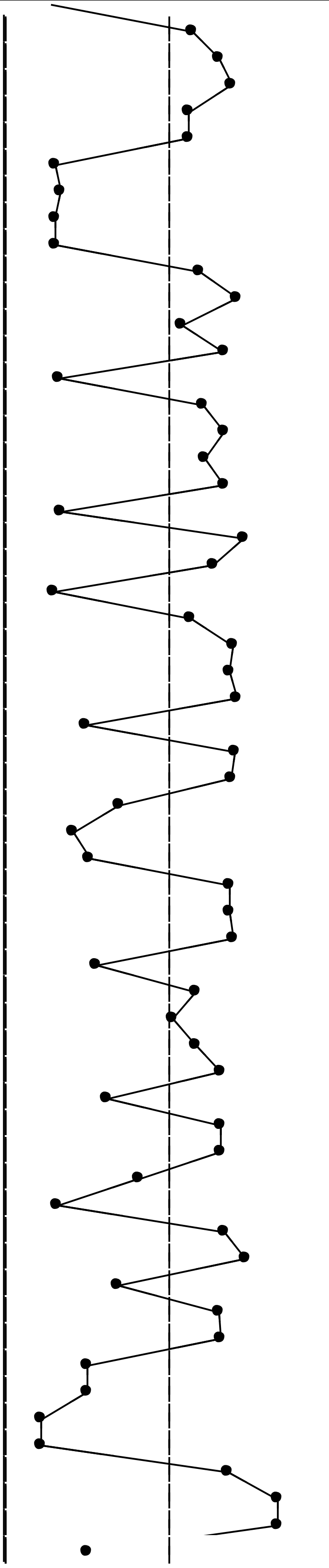
obs	Actual	Fitted	Residual
135499	1.00000	0.72085	0.27915
142439	1.00000	0.56632	0.43368
142550	0.00000	0.70737	-0.70737
142572	0.00000	0.42039	-0.42039
142586	1.00000	0.48695	0.51305
142666	0.00000	0.52297	-0.52297
142841	0.00000	0.55777	-0.55777
143061	1.00000	0.50282	0.49718
143641	1.00000	0.61452	0.38548
143673	1.00000	0.69971	0.30029
146338	1.00000	0.67161	0.32839
149285	0.00000	0.31423	-0.31423
255290	0.00000	0.85014	-0.85014
255454	0.00000	0.71813	-0.71813
255856	0.00000	0.70493	-0.70493
257138	0.00000	0.32137	-0.32137
257145	1.00000	0.56880	0.43120
264274	1.00000	0.73742	0.26258
264590	0.00000	0.60120	-0.60120
265104	0.00000	0.67366	-0.67366
265106	0.00000	0.67366	-0.67366
266646	1.00000	0.72294	0.27706
266661	0.00000	0.59832	-0.59832
268401	1.00000	0.48678	0.51322
270263	0.00000	0.59875	-0.59875
270541	1.00000	0.60219	0.39781
270772	1.00000	0.89360	0.10640
271001	1.00000	0.51857	0.48143
273059	0.00000	0.58174	-0.58174
273723	0.00000	0.34661	-0.34661
273971	0.00000	0.68779	-0.68779
275221	0.00000	0.42574	-0.42574
275367	1.00000	0.64805	0.35195
275368	1.00000	0.64805	0.35195
275464	1.00000	0.82590	0.17410
275512	1.00000	0.66584	0.33416
275513	1.00000	0.66584	0.33416
275579	0.00000	0.07129	-0.07129
275582	1.00000	0.81025	0.18975
275602	0.00000	0.07129	-0.07129
275830	1.00000	0.85787	0.14213
276353	0.00000	0.14249	-0.14249
276355	1.00000	0.85787	0.14213
276452	1.00000	0.27466	0.72534
276453	1.00000	0.27466	0.72534
276461	0.00000	0.57847	-0.57847
276462	0.00000	0.57847	-0.57847
276568	0.00000	0.75050	-0.75050

Residual Plot



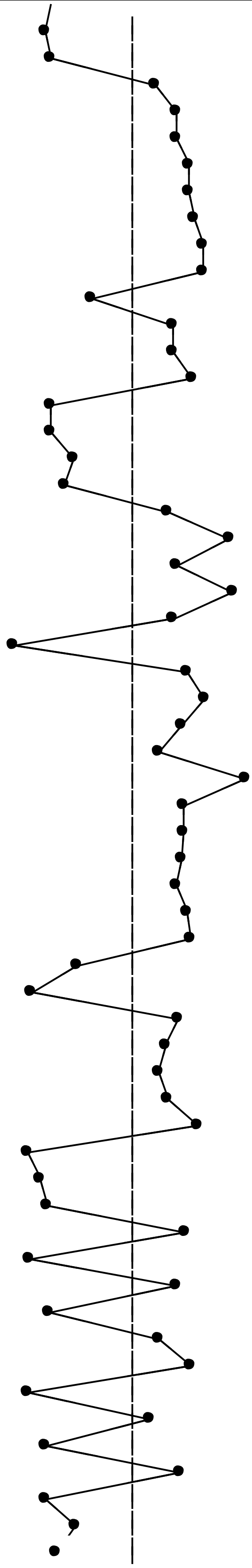
obs	Actual	Fitted	Residual
278231	1.00000	0.84673	0.15327
278253	1.00000	0.67134	0.32866
278333	1.00000	0.59212	0.40788
278419	1.00000	0.86935	0.13065
278420	1.00000	0.86935	0.13065
278468	0.00000	0.73215	-0.73215
278783	0.00000	0.68960	-0.68960
280598	0.00000	0.72802	-0.72802
280599	0.00000	0.72802	-0.72802
280830	1.00000	0.80195	0.19805
282252	1.00000	0.56481	0.43519
282567	1.00000	0.91128	0.08872
282597	1.00000	0.64860	0.35140
282602	0.00000	0.70255	-0.70255
282690	1.00000	0.77661	0.22339
282858	1.00000	0.64211	0.35789
282861	1.00000	0.76090	0.23910
282902	1.00000	0.64816	0.35184
282919	0.00000	0.69105	-0.69105
282927	1.00000	0.51591	0.48409
282939	1.00000	0.71057	0.28943
282957	0.00000	0.74274	-0.74274
283003	1.00000	0.86304	0.13696
283009	1.00000	0.58626	0.41374
283020	1.00000	0.60728	0.39272
283036	1.00000	0.55267	0.44733
283037	0.00000	0.53586	-0.53586
283089	1.00000	0.57318	0.42682
283093	1.00000	0.59659	0.40341
284499	0.00000	0.31965	-0.31965
284512	0.00000	0.61454	-0.61454
285249	0.00000	0.51181	-0.51181
285575	1.00000	0.60520	0.39480
285576	1.00000	0.60520	0.39480
287240	1.00000	0.57677	0.42323
287508	0.00000	0.47158	-0.47158
287943	1.00000	0.81524	0.18476
287963	1.00000	0.97095	0.02905
287994	1.00000	0.81524	0.18476
288036	1.00000	0.66159	0.33841
288388	0.00000	0.39287	-0.39287
288748	1.00000	0.66430	0.33570
288749	1.00000	0.66430	0.33570
288756	0.00000	0.18936	-0.18936
288811	0.00000	0.72247	-0.72247
288947	1.00000	0.63527	0.36473
289191	1.00000	0.49608	0.50392
289392	0.00000	0.32923	-0.32923
289746	1.00000	0.67852	0.32148
289947	1.00000	0.65946	0.34054
290203	0.00000	0.52208	-0.52208
290204	0.00000	0.52208	-0.52208
290298	0.00000	0.81582	-0.81582
290299	0.00000	0.81582	-0.81582
290431	1.00000	0.62313	0.37687
290665	1.00000	0.28902	0.71098
290666	1.00000	0.28902	0.71098
290781	0.00000	0.52833	-0.52833

Residual Plot



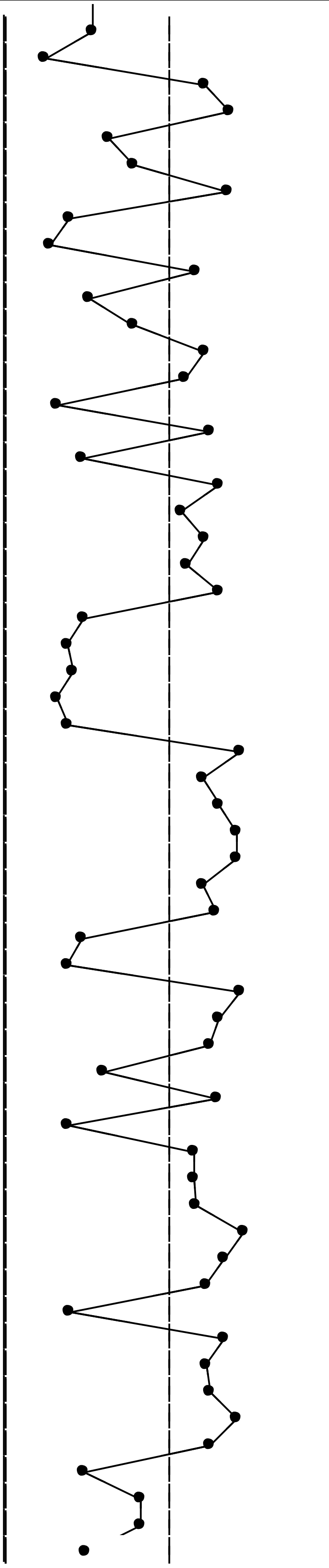
obs	Actual	Fitted	Residual
290829	0.00000	0.55351	-0.55351
291150	0.00000	0.52736	-0.52736
291185	1.00000	0.84065	0.15935
291264	1.00000	0.70091	0.29909
291265	1.00000	0.70091	0.29909
291778	1.00000	0.62725	0.37275
291779	1.00000	0.62725	0.37275
295236	1.00000	0.59212	0.40788
295364	1.00000	0.53988	0.46012
295365	1.00000	0.53988	0.46012
295677	0.00000	0.26515	-0.26515
295685	1.00000	0.72889	0.27111
295687	1.00000	0.72889	0.27111
295723	1.00000	0.60177	0.39823
295912	0.00000	0.52060	-0.52060
295913	0.00000	0.52060	-0.52060
296108	0.00000	0.36982	-0.36982
296220	0.00000	0.43125	-0.43125
296330	1.00000	0.76078	0.23922
296402	1.00000	0.36362	0.63638
296421	1.00000	0.71473	0.28527
296592	1.00000	0.33792	0.66208
296593	1.00000	0.72748	0.27252
296664	0.00000	0.76682	-0.76682
296682	1.00000	0.63885	0.36115
296688	1.00000	0.52532	0.47468
296745	1.00000	0.67852	0.32148
296798	1.00000	0.81639	0.18361
296818	1.00000	0.26288	0.73712
298446	1.00000	0.66112	0.33888
298613	1.00000	0.66071	0.33929
298785	1.00000	0.67483	0.32517
307067	1.00000	0.69933	0.30067
307246	1.00000	0.64743	0.35257
307314	1.00000	0.61411	0.38589
307428	0.00000	0.35116	-0.35116
307429	0.00000	0.65176	-0.65176
504125	1.00000	0.69059	0.30941
504268	1.00000	0.77905	0.22095
504299	1.00000	0.82360	0.17640
504458	1.00000	0.75995	0.24005
504524	1.00000	0.56809	0.43191
504579	0.00000	0.66521	-0.66521
504646	0.00000	0.59504	-0.59504
504656	0.00000	0.54779	-0.54779
505205	1.00000	0.65605	0.34395
505385	0.00000	0.66202	-0.66202
539616	1.00000	0.70088	0.29912
539959	0.00000	0.53869	-0.53869
671039	1.00000	0.81461	0.18539
671294	1.00000	0.60984	0.39016
671715	0.00000	0.68080	-0.68080
671882	1.00000	0.88609	0.11391
672681	0.00000	0.56197	-0.56197
673173	1.00000	0.68731	0.31269
673286	0.00000	0.55517	-0.55517
673898	0.00000	0.36647	-0.36647
675230	0.00000	0.49033	-0.49033

Residual Plot



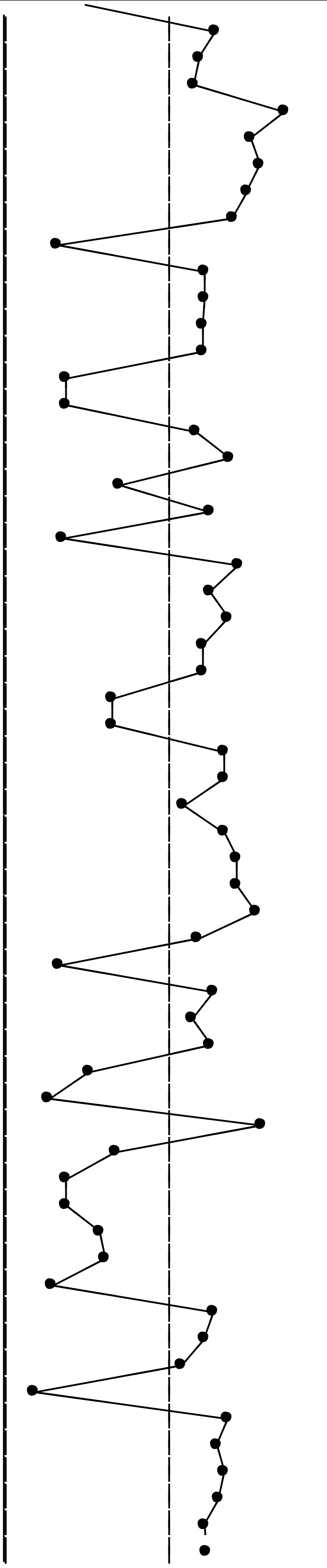
obs	Actual	Fitted	Residual
675231	0.00000	0.49033	-0.49033
676311	0.00000	0.80304	-0.80304
676450	1.00000	0.75945	0.24055
676533	1.00000	0.60266	0.39734
676604	0.00000	0.38334	-0.38334
676648	0.00000	0.22479	-0.22479
676659	1.00000	0.60984	0.39016
676756	0.00000	0.63203	-0.63203
676821	0.00000	0.76381	-0.76381
679535	1.00000	0.82644	0.17356
679545	0.00000	0.51252	-0.51252
679547	0.00000	0.22479	-0.22479
679556	1.00000	0.75945	0.24055
679778	1.00000	0.88855	0.11145
679821	0.00000	0.72081	-0.72081
681062	1.00000	0.73464	0.26536
681070	0.00000	0.56343	-0.56343
681225	1.00000	0.66653	0.33347
681254	1.00000	0.91494	0.08506
681285	1.00000	0.76303	0.23697
681290	1.00000	0.88609	0.11391
681441	1.00000	0.68030	0.31970
681774	0.00000	0.54706	-0.54706
681787	0.00000	0.65078	-0.65078
681990	0.00000	0.61180	-0.61180
682921	0.00000	0.71571	-0.71571
686002	0.00000	0.64129	-0.64129
686871	1.00000	0.54046	0.45954
686872	1.00000	0.77555	0.22445
686956	1.00000	0.67780	0.32220
687848	1.00000	0.55770	0.44230
687932	1.00000	0.55770	0.44230
688084	1.00000	0.77555	0.22445
688137	1.00000	0.69623	0.30377
688423	0.00000	0.56343	-0.56343
688431	0.00000	0.64129	-0.64129
688433	1.00000	0.54046	0.45954
688486	1.00000	0.68030	0.31970
688540	1.00000	0.73878	0.26122
688646	0.00000	0.42254	-0.42254
688650	1.00000	0.68943	0.31057
688653	0.00000	0.64885	-0.64885
688700	1.00000	0.83551	0.16449
688701	1.00000	0.83551	0.16449
688726	1.00000	0.82360	0.17640
688749	1.00000	0.51467	0.48533
688761	1.00000	0.64201	0.35799
688826	1.00000	0.75552	0.24448
688867	0.00000	0.63203	-0.63203
688868	1.00000	0.64632	0.35368
688874	1.00000	0.75552	0.24448
688994	1.00000	0.73462	0.26538
688997	1.00000	0.56632	0.43368
688999	1.00000	0.73739	0.26261
690032	0.00000	0.54779	-0.54779
690041	0.00000	0.18073	-0.18073
690042	0.00000	0.18073	-0.18073
690168	0.00000	0.54006	-0.54006

Residual Plot



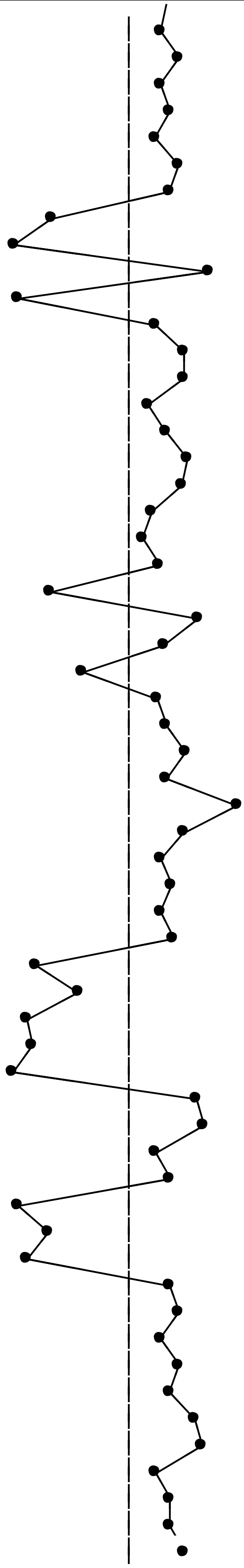
obs	Actual	Fitted	Residual
690450	1.00000	0.69623	0.30377
690724	1.00000	0.79865	0.20135
690812	1.00000	0.83300	0.16700
692015	1.00000	0.24914	0.75086
692039	1.00000	0.47439	0.52561
692398	1.00000	0.41175	0.58825
695164	1.00000	0.48695	0.51305
695356	1.00000	0.57861	0.42139
695403	0.00000	0.72081	-0.72081
695404	1.00000	0.76881	0.23119
695420	1.00000	0.75995	0.24005
695478	1.00000	0.77369	0.22631
695479	1.00000	0.77369	0.22631
695482	0.00000	0.65963	-0.65963
695483	0.00000	0.65963	-0.65963
695826	1.00000	0.82282	0.17718
695827	1.00000	0.59966	0.40034
697063	0.00000	0.32570	-0.32570
697269	1.00000	0.73034	0.26966
697543	0.00000	0.68544	-0.68544
697580	1.00000	0.54599	0.45401
697682	1.00000	0.73034	0.26966
697722	1.00000	0.61045	0.38955
698163	1.00000	0.77826	0.22174
698164	1.00000	0.77826	0.22174
698267	0.00000	0.36327	-0.36327
698268	0.00000	0.36327	-0.36327
698271	1.00000	0.64755	0.35245
698272	1.00000	0.64755	0.35245
698336	1.00000	0.90233	0.09767
698936	1.00000	0.64671	0.35329
698992	1.00000	0.55834	0.44166
698993	1.00000	0.55834	0.44166
729304	1.00000	0.42659	0.57341
729717	1.00000	0.80625	0.19375
741075	0.00000	0.71113	-0.71113
741618	1.00000	0.70592	0.29408
755323	1.00000	0.84572	0.15428
755724	1.00000	0.73878	0.26122
756231	0.00000	0.51492	-0.51492
756882	0.00000	0.77412	-0.77412
772377	1.00000	0.39897	0.60103
772692	0.00000	0.33609	-0.33609
775016	0.00000	0.65813	-0.65813
775017	0.00000	0.65813	-0.65813
775101	0.00000	0.44065	-0.44065
775255	0.00000	0.40440	-0.40440
775258	0.00000	0.74768	-0.74768
775280	1.00000	0.69860	0.30140
775299	1.00000	0.76303	0.23697
775308	1.00000	0.91494	0.08506
775352	0.00000	0.86492	-0.86492
775439	1.00000	0.60837	0.39163
775467	1.00000	0.68356	0.31644
775672	1.00000	0.63637	0.36363
779386	1.00000	0.67300	0.32700
779407	1.00000	0.76823	0.23177
866013	1.00000	0.75235	0.24765

Residual Plot



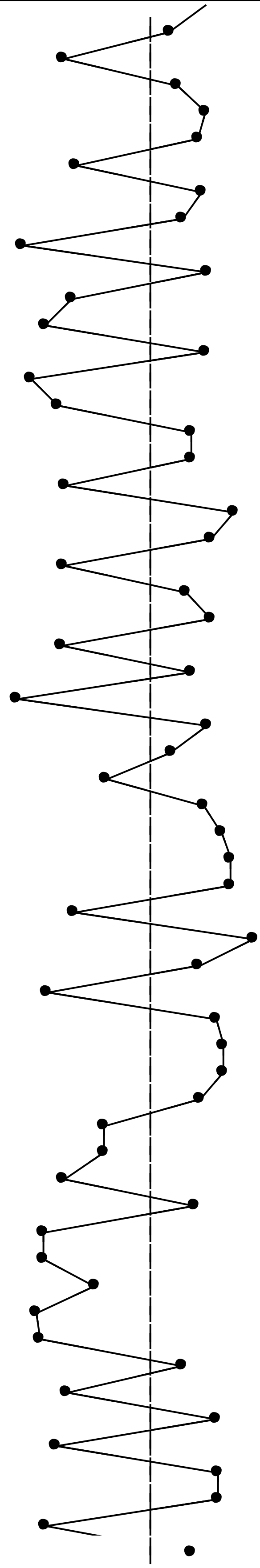
obs	Actual	Fitted	Residual
866034	1.00000	0.79110	0.20890
866035	1.00000	0.67592	0.32408
866038	1.00000	0.79520	0.20480
866050	1.00000	0.73562	0.26438
866054	1.00000	0.81874	0.18126
866056	1.00000	0.66677	0.33323
866060	1.00000	0.72467	0.27533
866097	0.00000	0.49028	-0.49028
866347	0.00000	0.72989	-0.72989
866641	1.00000	0.47885	0.52115
866652	0.00000	0.71398	-0.71398
866980	1.00000	0.81461	0.18539
870502	1.00000	0.63637	0.36363
870526	1.00000	0.63816	0.36184
870528	1.00000	0.86805	0.13195
870874	1.00000	0.75235	0.24765
870878	1.00000	0.61452	0.38548
870980	1.00000	0.65619	0.34381
876019	1.00000	0.84065	0.15935
876221	1.00000	0.90009	0.09991
876230	1.00000	0.80554	0.19446
876231	0.00000	0.49863	-0.49863
876246	1.00000	0.54565	0.45435
881098	1.00000	0.76881	0.23119
881832	0.00000	0.30049	-0.30049
882054	1.00000	0.80625	0.19375
882397	1.00000	0.75188	0.24812
882472	1.00000	0.62686	0.37314
882668	1.00000	0.75058	0.24942
885058	1.00000	0.28643	0.71357
888171	1.00000	0.64433	0.35567
888239	1.00000	0.78215	0.21785
888280	1.00000	0.72297	0.27703
888296	1.00000	0.78693	0.21307
888323	1.00000	0.71077	0.28923
888451	0.00000	0.59085	-0.59085
888760	0.00000	0.31812	-0.31812
888801	0.00000	0.64130	-0.64130
888805	0.00000	0.61786	-0.61786
888806	0.00000	0.74211	-0.74211
888849	1.00000	0.56363	0.43637
896673	1.00000	0.51467	0.48533
896674	1.00000	0.82644	0.17356
897339	1.00000	0.73406	0.26594
897366	0.00000	0.70415	-0.70415
899070	0.00000	0.51252	-0.51252
899187	0.00000	0.64885	-0.64885
899951	1.00000	0.73742	0.26258
902191	1.00000	0.66677	0.33323
904881	1.00000	0.79110	0.20890
905243	1.00000	0.67592	0.32408
916235	1.00000	0.72467	0.27533
916658	1.00000	0.57135	0.42865
922053	1.00000	0.52808	0.47192
922819	1.00000	0.82590	0.17410
923551	1.00000	0.73562	0.26438
929021	1.00000	0.73739	0.26261
929024	1.00000	0.63816	0.36184

Residual Plot



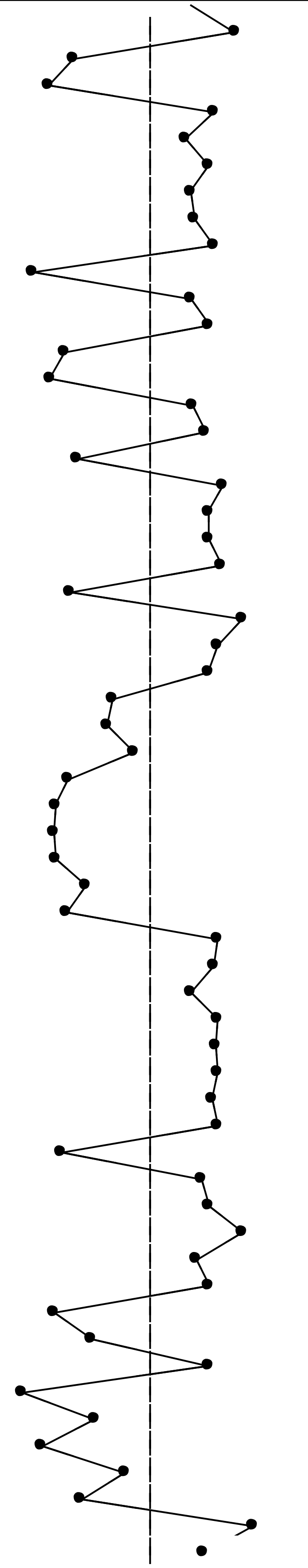
obs	Actual	Fitted	Residual
929048	1.00000	0.86805	0.13195
929054	0.00000	0.56331	-0.56331
929057	1.00000	0.81874	0.18126
929060	1.00000	0.64632	0.35368
929100	1.00000	0.68943	0.31057
929118	0.00000	0.48240	-0.48240
929129	1.00000	0.65973	0.34027
929228	1.00000	0.78887	0.21113
929242	0.00000	0.81521	-0.81521
929267	1.00000	0.63126	0.36874
929273	0.00000	0.50222	-0.50222
933063	0.00000	0.67671	-0.67671
936292	1.00000	0.64671	0.35329
936644	0.00000	0.76143	-0.76143
936913	0.00000	0.59542	-0.59542
944429	1.00000	0.73462	0.26538
951237	1.00000	0.73469	0.26531
951473	0.00000	0.54541	-0.54541
951755	1.00000	0.45676	0.54324
951778	1.00000	0.60177	0.39823
952264	0.00000	0.56173	-0.56173
952272	1.00000	0.75970	0.24030
952285	1.00000	0.60219	0.39781
982099	0.00000	0.56661	-0.56661
992594	1.00000	0.72759	0.27241
993501	0.00000	0.86188	-0.86188
997966	1.00000	0.62715	0.37285
13254P	1.00000	0.86125	0.13875
13381M	0.00000	0.27804	-0.27804
13395E	1.00000	0.65605	0.34395
13396U	1.00000	0.54018	0.45982
13410M	1.00000	0.48326	0.51674
13410N	1.00000	0.48326	0.51674
13518D	0.00000	0.48829	-0.48829
13561E	1.00000	0.33400	0.66600
13585K	1.00000	0.68208	0.31792
13598C	0.00000	0.66335	-0.66335
13611T	1.00000	0.56848	0.43152
13635T	1.00000	0.52149	0.47851
13635V	1.00000	0.52149	0.47851
13653D	1.00000	0.66786	0.33214
13703L	0.00000	0.29428	-0.29428
13703M	0.00000	0.29428	-0.29428
13787Q	0.00000	0.55419	-0.55419
13831Q	1.00000	0.71442	0.28558
13922L	0.00000	0.68949	-0.68949
13922N	0.00000	0.68949	-0.68949
13927H	0.00000	0.34334	-0.34334
13928H	0.00000	0.72591	-0.72591
14008E	0.00000	0.71392	-0.71392
14051H	1.00000	0.79052	0.20948
14308E	0.00000	0.53667	-0.53667
14308F	1.00000	0.57460	0.42540
14318F	0.00000	0.60198	-0.60198
14360D	1.00000	0.56214	0.43786
14360E	1.00000	0.56214	0.43786
14402K	0.00000	0.67855	-0.67855
14479T	1.00000	0.73271	0.26729

Residual Plot



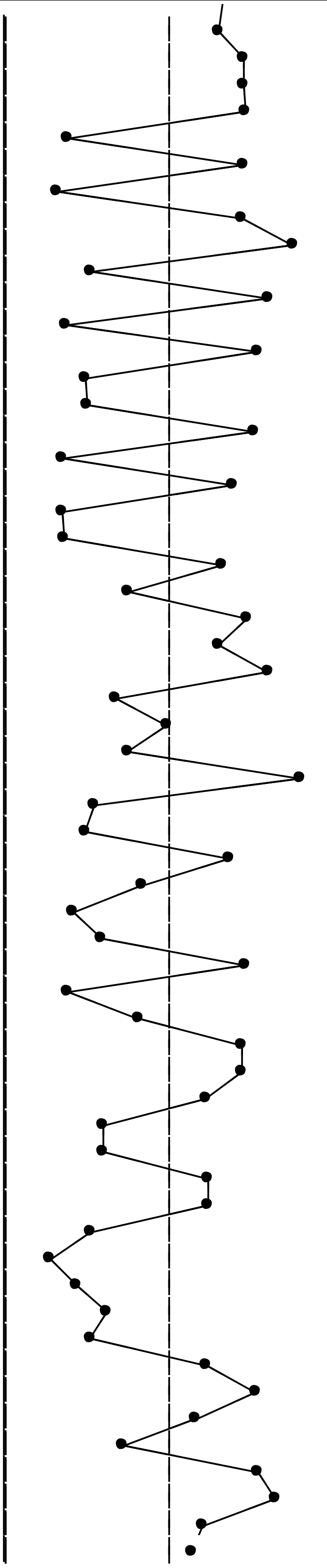
obs	Actual	Fitted	Residual
14483C	1.00000	0.45005	0.54995
14489M	0.00000	0.48333	-0.48333
14494T	0.00000	0.64130	-0.64130
14563F	1.00000	0.57648	0.42352
14572M	1.00000	0.76371	0.23629
14581K	1.00000	0.62163	0.37837
14583R	1.00000	0.73021	0.26979
14591U	1.00000	0.71442	0.28558
14612H	1.00000	0.58467	0.41533
14628F	0.00000	0.75346	-0.75346
14638C	1.00000	0.73741	0.26259
14648T	1.00000	0.61794	0.38206
14750V	0.00000	0.54834	-0.54834
14807P	0.00000	0.63836	-0.63836
15187C	1.00000	0.71792	0.28208
15224L	1.00000	0.64674	0.35326
15297U	0.00000	0.45842	-0.45842
15300L	1.00000	0.52808	0.47192
15303E	1.00000	0.60886	0.39114
15304M	1.00000	0.62301	0.37699
15304N	1.00000	0.53441	0.46559
15305L	0.00000	0.51010	-0.51010
15305R	1.00000	0.39786	0.60214
15305T	1.00000	0.55583	0.44417
15307J	1.00000	0.61664	0.38336
15307N	0.00000	0.22792	-0.22792
15307Q	0.00000	0.27585	-0.27585
15311F	0.00000	0.10062	-0.10062
15312E	0.00000	0.52786	-0.52786
15312F	0.00000	0.60021	-0.60021
15316E	0.00000	0.61832	-0.61832
15316F	0.00000	0.60422	-0.60422
15316V	0.00000	0.40648	-0.40648
15318F	0.00000	0.53596	-0.53596
15318N	1.00000	0.55134	0.44866
15318U	1.00000	0.57900	0.42100
15435F	1.00000	0.73271	0.26729
15445U	1.00000	0.55834	0.44166
25728F	1.00000	0.56715	0.43285
26301U	1.00000	0.55859	0.44141
26469J	1.00000	0.59589	0.40411
27221L	1.00000	0.55822	0.44178
27253F	0.00000	0.57032	-0.57032
27414C	1.00000	0.66371	0.33629
27879H	1.00000	0.62104	0.37896
27948F	1.00000	0.39627	0.60373
28260W	1.00000	0.69393	0.30607
28362V	1.00000	0.62380	0.37620
28444N	0.00000	0.61588	-0.61588
28444Q	0.00000	0.37465	-0.37465
28467J	1.00000	0.60945	0.39055
28638E	0.00000	0.82495	-0.82495
28643K	0.00000	0.35354	-0.35354
28667X	0.00000	0.69761	-0.69761
28672F	0.00000	0.15850	-0.15850
28916J	0.00000	0.44847	-0.44847
28916T	1.00000	0.33481	0.66519
28982C	1.00000	0.65610	0.34390

Residual Plot



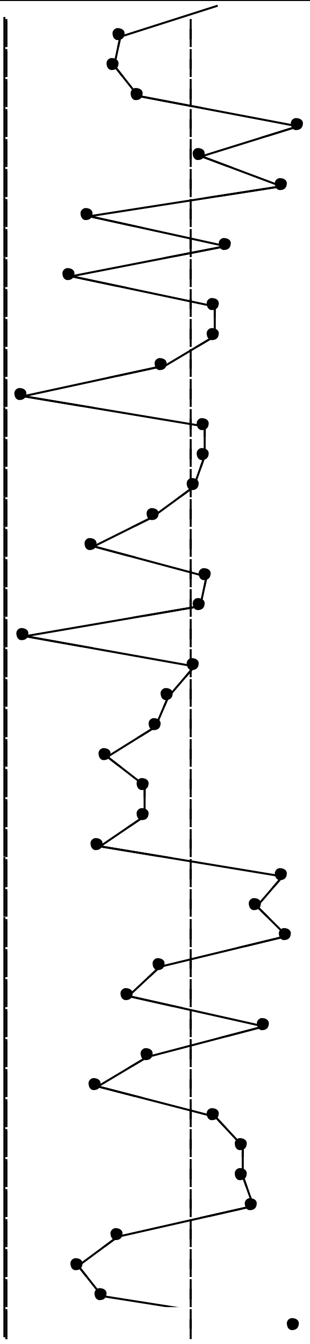
obs	Actual	Fitted	Residual
28985F	1.00000	0.67475	0.32525
28990K	1.00000	0.51121	0.48879
28990N	1.00000	0.51121	0.48879
29069V	1.00000	0.49824	0.50176
29082M	0.00000	0.64407	-0.64407
29209T	1.00000	0.51641	0.48359
29241J	0.00000	0.71866	-0.71866
29280M	1.00000	0.53046	0.46954
29530P	1.00000	0.18832	0.81168
29567H	0.00000	0.50334	-0.50334
29676X	1.00000	0.35509	0.64491
29927C	0.00000	0.66216	-0.66216
29985K	1.00000	0.42281	0.57719
29985U	0.00000	0.53918	-0.53918
30166D	0.00000	0.52660	-0.52660
30364F	1.00000	0.44232	0.55768
30424E	0.00000	0.68249	-0.68249
30449D	1.00000	0.57815	0.42185
30896V	0.00000	0.68173	-0.68173
30911D	0.00000	0.66527	-0.66527
30918X	1.00000	0.65190	0.34810
30926V	0.00000	0.25289	-0.25289
30985R	1.00000	0.48493	0.51507
31131W	1.00000	0.66557	0.33443
31145C	1.00000	0.35521	0.64479
31233L	0.00000	0.33709	-0.33709
31451L	0.00000	0.00843	-0.00843
31482P	0.00000	0.26715	-0.26715
31487L	1.00000	0.14359	0.85641
31539J	0.00000	0.48221	-0.48221
31555N	0.00000	0.54058	-0.54058
31674U	1.00000	0.60761	0.39239
31860R	0.00000	0.16358	-0.16358
31924K	0.00000	0.60722	-0.60722
31933H	0.00000	0.42665	-0.42665
31964P	1.00000	0.50173	0.49827
32001T	0.00000	0.65452	-0.65452
32096F	0.00000	0.18672	-0.18672
32146D	1.00000	0.52760	0.47240
32146E	1.00000	0.52760	0.47240
32269V	1.00000	0.75131	0.24869
32331J	0.00000	0.42230	-0.42230
32335C	0.00000	0.42230	-0.42230
32374F	1.00000	0.74102	0.25898
32374W	1.00000	0.74102	0.25898
32503H	0.00000	0.50612	-0.50612
32523D	0.00000	0.75852	-0.75852
32526T	0.00000	0.59493	-0.59493
32565E	0.00000	0.39810	-0.39810
32585H	0.00000	0.50084	-0.50084
32616V	1.00000	0.74948	0.25052
32667F	1.00000	0.43279	0.56721
35612E	1.00000	0.82275	0.17725
35738F	0.00000	0.28697	-0.28697
36066K	1.00000	0.42502	0.57498
36167F	1.00000	0.30411	0.69589
36230C	1.00000	0.77999	0.22001
41357X	1.00000	0.85285	0.14715

Residual Plot



obs	Actual	Fitted	Residual
50343P	0.00000	0.39717	-0.39717
50397K	0.00000	0.43043	-0.43043
51272T	0.00000	0.29468	-0.29468
69171K	1.00000	0.37653	0.62347
69238T	1.00000	0.94299	0.05701
74142F	1.00000	0.47527	0.52473
74166N	0.00000	0.57958	-0.57958
74653M	1.00000	0.79690	0.20310
75672V	0.00000	0.68400	-0.68400
75678H	1.00000	0.85597	0.14403
75687X	1.00000	0.85597	0.14403
75729P	0.00000	0.15854	-0.15854
77204T	0.00000	0.96205	-0.96205
77287U	1.00000	0.91459	0.08541
77287V	1.00000	0.91459	0.08541
77315P	1.00000	0.97091	0.02909
8651ZM	0.00000	0.20700	-0.20700
8680LG	0.00000	0.55572	-0.55572
86829U	1.00000	0.90014	0.09986
86829V	1.00000	0.93096	0.06904
86987M	0.00000	0.94829	-0.94829
86996Q	1.00000	0.97263	0.02737
87279Q	0.00000	0.12453	-0.12453
87399C	0.00000	0.18741	-0.18741
8750JA	0.00000	0.47902	-0.47902
87521U	0.00000	0.25452	-0.25452
87673F	0.00000	0.25648	-0.25648
88214M	0.00000	0.52663	-0.52663
88221W	1.00000	0.46453	0.53547
88229C	1.00000	0.60856	0.39144
88311E	1.00000	0.43575	0.56425
88330H	0.00000	0.16641	-0.16641
88390U	0.00000	0.34867	-0.34867
8848TW	1.00000	0.57135	0.42865
88864M	0.00000	0.23724	-0.23724
89259R	0.00000	0.54049	-0.54049
89280N	1.00000	0.86237	0.13763
89368N	1.00000	0.69598	0.30402
89387E	1.00000	0.69598	0.30402
8942D7	1.00000	0.64345	0.35655
91291R	0.00000	0.41364	-0.41364
93757Y	0.00000	0.63886	-0.63886
9506RZ	0.00000	0.50222	-0.50222
9514X1	1.00000	0.39826	0.60174

Residual Plot



Source: Author’s own elaboration.

Appendix 3 - Goodness-of-fit evaluation for binary specification (Andrews and Hosmer-Lemeshow Tests)

	Quantile of Risk		RD = 0		RD = 1		Total	H-L
	Low	High	Actual	Expect	Actual	Expect	Obs	Value
1	0,0084	0,3466	42	41,6069	13	13,3931	55	0,0153
2	0,3487	0,4868	31	32,3886	25	23,6114	56	0,1412
3	0,4869	0,5401	31	26,5943	24	28,4057	55	1,4132
4	0,5402	0,5796	22	24,6017	34	31,3983	56	0,4907
5	0,5817	0,6320	18	21,9410	38	34,0590	56	1,1639
6	0,6320	0,6643	20	19,2692	35	35,7308	55	0,0427
7	0,6652	0,7049	20	17,7168	36	38,2832	56	0,4304
8	0,7059	0,7477	18	14,9960	37	40,0040	55	0,8274
9	0,7495	0,8152	9	12,5315	47	43,4685	56	1,2821
10	0,8152	0,9726	8	7,3542	48	48,6458	56	0,0653
		Total	219	219	337	337	556	5,8721
H-L Statistic				5,8720	Prob. Chi-Sq(8)			0,6616
Andrews Statistic				6,2098	Prob. Chi-Sq(10)			0,7973

Source: Author’s own elaboration.

8. References

Aghion, P., *et al.* (1998), *Endogenous growth theory.*

Allison, P. D. (2014), Measures of fit for logistic regression

Arora, A., *et al.* (2008), "R&D and the patent premium", *International Journal of Industrial Organization,* Vol. 26, No. 5, pp. 1153-1179.

Arora, A. and A. Gambardella (1990), "Complementarity and External Linkages: The Strategies of the Large Firms in Biotechnology", *The Journal of Industrial Economics,* Vol. 38, No. 4, pp. 361-379.

Bhagat, S. and I. Welch (1995), "Corporate research & development investments international comparisons", *Journal of Accounting and Economics,* Vol. 19, No. 2–3, pp. 443-470.

Braga, H. and L. Willmore (1991), "Technological Imports and Technological Effort: An Analysis of their Determinants in Brazilian Firms", *The Journal of Industrial Economics,* Vol. 39, No. 4, pp. 421-432.

Chen, W. R. and K. D. Miller (2007), "Situational and institutional determinants of firms' R&D search intensity", *Strategic Management Journal,* Vol. 28, No. 4, pp. 369-381.

Coad, A. and R. Rao (2010), "Firm growth and R&D expenditure", *Economics of Innovation and New Technology,* Vol. 19, No. 2, pp. 127-145.

Cohen, W. M. and S. Klepper (1996), "A Reprise of Size and R & D", *The Economic Journal,* Vol. 106, No. 437, pp. 925-951.

Cohen, W. M. and R. C. Levin (1989), "Empirical studies of innovation and market structure", *Handbook of Industrial Organization,* Vol. No. pp. 1060-1107.

Cohen, W. M., *et al.* (1987), "Firm size and R&D intensity: A re-examination", *Journal of Industrial Economics,* Vol. 35, No. 1987, pp. 543-565.

Costa-Campi, M. T., *et al.* (2014), "R&D drivers and obstacles to innovation in the energy industry", *Energy Economics,* Vol. 46, No. 2014, pp. 20-30.

Evans, G., *et al.* (1996), "Measuring Left-Right and Libertarian-Authoritarian Values in the British Electorate", *The British Journal of Sociology,* Vol. 47, No. 1, pp. 93-112.

Fishman, A. and R. Rob (1999), "The size of firms and R&D investment", *International Economic Review,* Vol. 40, No. 4, pp. 915-931.

Freeman, C. (1991), "Networks of innovators: A synthesis of research issues", *Research Policy,* Vol. 20, No. 5, pp. 499-514.

Galende Del Canto, J. and I. Suárez González (1999), "A resource-based analysis of the factors determining a firm's R&D activities", *Research Policy,* Vol. 28, No. 8, pp. 891-905.

Galende, J. and J. M. Fuente (2003), "Internal factors determining a firm’s innovative behaviour", *Research Policy,* Vol. 32, No. 5, pp. 715-736.

Ginarte, J. C. and W. G. Park (1997), "Determinants of patent rights: A cross-national study", *Research Policy,* Vol. 26, No. 3, pp. 283-301.

Griffith, R.*, et al.* (2006), "Innovation and Productivity Across Four European Countries", *Oxford Review of Economic Policy,* Vol. 22, No. 4, pp. 483-498.

Guevara, H. H.*, et al.* (2015), The 2015 EU Industrial R&D Investment Scoreboard Institute for Prospective and Technological Studies, Joint Research Centre

Gujarati, D. N. (2009), *Basic econometrics.*

Gustavsson, P. and A. Poldahl (2003), *Determinants of firm R&D: Evidence from Swedish firm level data.*

Hall, B. H. (2010), "The financing of innovative firms", *Review of Economics and Institutions,* Vol. 1, No. 1, pp.

Huang, C.*, et al.* (2015), *Emerging Economies, Risk and Development, and Intelligent Technology: Proceedings of the 5th International Conference on Risk Analysis and Crisis Response, June 1-3, 2015, Tangier, Morocco.*

Hunt, S. D. and R. M. Morgan (1996), "The Resource-Advantage Theory of Competition: Dynamics, Path Dependencies, and Evolutionary Dimensions", *Journal of Marketing,* Vol. 60, No. 4, pp. 107-114.

Kamien, M. I. and N. L. Schwartz (1975), "Market Structure and Innovation: A Survey", *Journal of Economic Literature,* Vol. 13, No. 1, pp. 1-37.

Kim, H. and S.-Y. Park (2012), "THE RELATION BETWEEN CASH HOLDINGS AND R&D EXPENDITURES ACCORDING TO OWNERSHIP STRUCTURE", *Eurasian Business Review,* Vol. 2, No. 2, pp. 25-42.

Lai, Y.-L.*, et al.* (2015), "Factors affecting firm's R&D investment decisions", *Journal of Business Research,* Vol. 68, No. 4, pp. 840-844.

Limanlı, Ö. (2015), "Determinants of R&D Investment Decision in Turkey", *Procedia - Social and Behavioral Sciences,* Vol. 195, No. 2015, pp. 759-767.

Long, J. S. and J. Freese (2006), *Regression models for categorical dependent variables using Stata.*

Marr, B. (2008), Impacting future value: how to manage your intellectual capital CMA Canada Mississauga, ON

McCartney, M. (2006), "Can a heterodox economist use cross-country growth regressions?", *Post-autistic economics review,* Vol. 37, No. pp. 45-54.

Myers, S. C. (1984), "The Capital Structure Puzzle", *The Journal of Finance,* Vol. 39, No. 3, pp. 574-592.

OECD (2002), *Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development.*

OECD (2005), "Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data", Vol. No. 3rd Edition, pp. 166.

Park, H.-A. (2013), "An Introduction to Logistic Regression: From Basic Concepts to Interpretation with Particular Attention to Nursing Domain", *J Korean Acad Nurs,* Vol. 43, No. 2, pp. 154-164.

Reynard, E. (1979), "METHOD FOR RELATING RESEARCH SPENDING TO NET PROFITS", *Research Management,* Vol. 22, No. 4, pp. 12-14.

Romer, P. M. (1994), "The origins of endogenous growth", *The journal of economic perspectives,* Vol. 8, No. 1, pp. 3-22.

Schumpeter, J. A. (1934), *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle.* Cambridge, MA: H. U. Press

Schumpeter, J. A. (1939), *Business Cycles: a Theoretical, Historical and Statistical Analysis of the Capitalist Process.* New York: McGraw-Hill

Schumpeter, J. A. (1942), *Capitalism, socialism and democracy.* New York: G. A. U. L. r. 1976)

Siddharthan, N. S. (1992), "Transaction costs, technology transfer, and in-house R&D", *Journal of Economic Behavior & Organization,* Vol. 18, No. 2, pp. 265-271.

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations.*

Solow, R. M. (1956), "A Contribution to the Theory of Economic Growth", *The Quarterly Journal of Economics,* Vol. 70, No. 1, pp. 65-94.

Standfield, K. (2005), *Intangible finance standards: Advances in fundamental analysis and technical analysis.*

Teece, D. J. (1986), "Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy", *Research Policy,* Vol. 15, No. 6, pp. 285-305.

Urbancová, H. (2013), "Competitive Advantage Achievement through Innovation and Knowledge", *Journal of Competitiveness*, Vol. 5, No. 1, pp. 82-96.

Varsakelis, N. C. (2001), "The impact of patent protection, economy openness and national culture on R&D investment: a cross-country empirical investigation", *Research Policy,* Vol. 30, No. 7, pp. 1059-1068.

Wang, E. C. (2010), "Determinants of R&D investment: The Extreme-Bounds-Analysis approach applied to 26 OECD countries", *Research Policy,* Vol. 39, No. 1, pp. 103-116.